

Thought Experiments between Nature and Society

Thought Experiments between Nature and Society:

A Festschrift for Nenad Mišćević

Edited by

Bojan Borstner and Smiljana Gartner

Cambridge
Scholars
Publishing



Thought Experiments between Nature and Society:
A Festschrift for Nenad Mišćević

Edited by Bojan Borstner and Smiljana Gartner

This book first published 2017

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Copyright © 2017 by Bojan Borstner, Smiljana Gartner and contributors

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-4438-8643-2

ISBN (13): 978-1-4438-8643-7

TABLE OF CONTENTS

Preface	ix
Introduction: Thought Experiments as an Important Part of Philosophical Methodology	xiii
Contributors	xxx
Imagining Nenad	1
Bojan Borstner, Tadej Todorović	
Accounting for Thought Experiments – 25 Years Later.....	11
Nenad Mišćević	
Part I: Thought Experiment – General	
From Anti-Metaphysics to Metaphysics.....	32
Timothy Williamson	
Nenad’s Reply	45
Intuitions and Thought Experiments	49
Howard Robinson	
Nenad’s Reply	62
Mišćević on Intuitions and Thought Experiments	66
Maja Malec, Olga Markić	
Nenad’s Reply	79
Curiosity and the Argumentative Process.....	83
Nenad Smokrović	
Nenad’s Reply	96

Semantic Transformations	102
Peter Gärdenfors	
Nenad's Reply	115
Lucky Math: Anti-luck Epistemology and Necessary Truth	119
Danilo Šuster	
Nenad's Reply	134
Epistemic Luck and Logical Necessities: Armchair Luck Revisited.....	137
Guido Melchior	
Nenad's Reply	151
Did a Particularist Kill the Thought Experiments?	154
Smiljana Gartner	
Nenad's Reply	166
Part II: Thought Experiments and Nature	
Experimental Philosophy, Gettier-Cases, and Pragmatic Projection	172
Marian David	
Nenad's Reply	185
Concepts in a World of Particulars	189
Peter Simons	
Nenad's Reply	200
Is the Speed of Light Knowable A Priori?.....	204
Ilhan Inan	
Nenad's Reply	216
Mental Models in Scientific Work	219
Andrej Ule	
Nenad's Reply	234

Natural Kinds and Conceptual Truth	239
Ferenc Huoranszki	
Nenad's Reply	252
Grasping the Basic Arithmetical Concepts: The Role of Imaginative Intuitions.....	257
Majda Trobok	
Nenad's Reply	269
The Colour Dilemma: A Subjectivist Answer	274
Andraž Stožer, Janez Bregant	
Nenad's Reply	286
Dasein in a Vat	291
Matjaž Potrč	
Nenad's Reply	297
Part III: Thought Experiments and Society	
Knowing One's Own Mind	302
Pierre Jacob	
Nenad's Reply	323
The False Promise of Thought-Experimentation in Moral and Political Philosophy	328
Friderik Klampfer	
Nenad's Reply	349
Miščević, Mental Models, and Thought Experiments in Political Philosophy	354
Miomir Matulović	
Nenad's Reply	370

Are Nations Social Constructs? Nenad on Nations	375
Boran Berčić	
Nenad's Reply	386
Thought Experiments in Teaching: TE as a Suppositional Real Story	390
Rudi Kotnik	
Nenad's Reply	406
The Ring of Gyges and the Philosophical Imagination	410
Boris Vežjak	
Nenad's Reply	425
Index	430

PREFACE

“Suppose now that there were two such magic rings, and the just put on one of them and the unjust the other; no man can be imagined (hos doxeien) to be of such an iron nature that he would stand fast in justice...”
(Plato)

“Indeed, the possibility of thought experiments rests upon our ideas as being the more or less exact copy of facts.” (Ernst Mach)

“I was sitting on a chair in my patent office in Bern. Suddenly a thought struck me: If a man falls freely, he would not feel his weight. I was taken aback. This simple thought experiment made a deep impression on me. This led me to the theory of gravity.” (Albert Einstein)

The idea for this book has its roots in a discussion that happened in the middle of May in 2014, when the researchers of the Department of Philosophy, Faculty of Arts, University of Maribor, were completing the work on the research program *The Concept of Virtue in Theoretical and Practical Philosophy*. We asked ourselves: “What should be the main subject of our future research program?” We were operating with the idea that contemporary philosophy is often very fragmented, with extremely narrow specialization. Having this in mind, we came, after a long discussion, to a proposal which should be able to rectify this narrowness – thought experiment as a common place or starting point that can serve as a common frame for our research in different fields of philosophy. Our proposal for the research framework has been approved, and starting in 2015 we began our research work in the research program *Laboratorij duha: Miselni eksperimenti med naravo in družbo (Laboratory of Mind: Thought Experiments from Nature and Society)*, with Professor Miščević being the leading researcher. There is another additional point worth mentioning: in 2015, Professor Miščević celebrated his sixty-fifth birthday, so combining the research task with his celebration seemed like the obvious thing to do. Thus we also organized the conference in November 2015 with the same title as this book has.

What are thought experiments (TE), and why are TEs something that can help us in the defragmentation of, at least in a methodological sense, contemporary philosophy? There are many possible ways to answer these

questions. We will highlight only one: just start with some paradigm cases and, through analysis, step by step come to the most important properties, features of TEs. We have been told that in philosophy and science one encounters a lot of TEs; for instance, the story of the *Ring of Gyges* in the second book of Plato's *Republic*. Indeed, the entire *Republic* might be viewed as an extended TE, in which the picture of an ideal state is built "in the logos", as Plato puts it. Is this really a TE? There is an ongoing debate between those who try to justify the idea that even some of the Presocratic philosophers of nature provide us with a bunch of good examples of TEs and others who claim just the opposite and seriously doubt the possibility of using the term as referring to the ancient practice of appealing to intuition. However, no one doubts that in TEs, if there are such cases, there is a special kind of reasoning at work which is often described as 'armchair' reasoning, in contrast to laboratory or outdoor research. It is not only used in philosophy but also in science (e.g. physics, see famous TEs by Einstein and Galileo, economics, etc.). In philosophy, it is used in political philosophy, ethics, ontology, philosophy of mind, and epistemology. Philosophers reflect on such knowledge, and in the last twenty-five years, they have gained support from cognitive scientists in this effort. The cooperation between cognitive scientists and philosophers offers a chance for a unified, respectably scientific account of human knowledge.

When Ernest Mach starts to think about TEs (Gedankenexperimente), he realizes that, in a certain sense, any sensibly designed real experiment should be preceded by a thought experiment which anticipates at any rate the possibility of its outcome. We can add that the purpose of a TE is to strip away the things that complicate matters in reality in order to focus more clearly on the most important properties of the problem that we are trying to solve.

Typical stages in thought experimenting and reflection are understanding, building a model, reaching a particular judgment-intuition, 'intuitive induction', and then a leap to a more general judgment. However, there are at least three points which are closely connected with our understanding of TEs.

Firstly, explanatory matters. Are TEs reliable sources of data for philosophers and scientists? If no, how should philosophy proceed? If yes, how is their reliability to be accounted for? The mysterious Platonic 'perception' does not sound like a good explanation, but the opposite stance, that judgments-intuitions from TEs are just ordinary judgments,

seems an oversimplification. We need an explanatory proposal, and a TE tries to offer one.

Secondly, normative problem – a priori and a posteriori. In a typical TE one can separate the contribution of general empirical knowledge from topic-specific (mathematical, moral, and probably normative epistemological, etc.) contribution, which might be deeply or superficially a priori. Therefore, their typical products, intuitional judgments, have a structured justification encompassing a priori and a posteriori elements. Most of them are *prima facie* justified without appeal to experience, but only *prima facie*. Full reflective justification of a given intuition is a posteriori. For instance, in the case of kind-intuitions, a posteriori justification derives from the general stability and success of thinking in terms of kinds. In the case of knowledge, from the general stability and success of appealing to knowledge, and distinguishing it from its lesser cousins.

Thirdly, application issues. How could TE-intuitions be made better? How could people, in principle, come to ameliorate their intuitions? How could enlightened individuals (experts) guide novices to avoid typical mistakes? TEs have stages, including reflection on them. Therefore, look at the stages; divide labour among stages – it might help. Dangers and promises are unequally distributed in the biography of a typical philosophical intuition – dangers abound in early stages, promises come in later. Collating TEs and performing intuitive inductions might help (specially to counter the opacity of each intuition taken in isolation); individual reflective equilibrium as well. It caters for coherence, in good cases. One then recognizes the margins of one's own reliability of intuiting. Finally, the collective wide reflexive equilibrium brings in factual knowledge; it should help. So, in principle, we can (hopefully) avoid the problem of regress or circularity, and show how expertize is possible.

To summarize, the basic goal of this book is to deepen the understanding of TEs. The additional goal is defending their reliability against sceptics. The practical goal is proposing new applications of TEs, both in theoretical philosophy and in ethical and political thinking, with important practical consequences.

And the last but not least important goal of the book is to honour Professor Mišćević at his 65th birthday.

We received invaluable research assistance from Tadej Todorović in the final stages of the project.

Financial support for the project as a whole was provided by the Slovenian Research Agency (research core funding No. P6-0144).

Bojan Borstner, Smiljana Gartner

INTRODUCTION: THOUGHT EXPERIMENTS AS AN IMPORTANT PART OF PHILOSOPHICAL METHODOLOGY

BOJAN BORSTNER, SMILJANA GARTNER
UNIVERSITY OF MARIBOR

Thought experiments (TEs for short) have become a hot topic in the methodology of philosophy as well as in philosophy of science.

The reason for that is that they have been present throughout the whole history of the discipline, and have become indispensable for analytic philosophy. Collections of papers dedicated to particular philosophical topics or arguments are nowadays often organized around famous TEs, say Brain-in-a-Vat, or Original Position, or Twin Earth and its cognates. The indispensability is at least practical and valid for normative disciplines, from epistemology, through philosophy of language, to ethics, and political philosophy. In descriptive-explanatory disciplines like metaphysics they are required for arriving at a good pattern of our everyday concepts (for instance, identity of a material object through time) to be compared with the counterpart scientific one.

Let us start with a few examples of famous thought experiments. For historical examples think of Plato's Glaucon (*The Republic*: 359-360):

Glaucon: But as for the second point, that those who practice it do so unwillingly and from want of power to commit injustice—we shall be most likely to apprehend that if we entertain some such supposition as this in thought: *tei dianoi*ai if we grant to each, the just and the unjust, licence and power to do whatever he pleases, and then accompany them in imagination and see whither his desire will conduct each.

Suppose now that there were two such magic rings, and the just put on one of them and the unjust the other; no man can be imagined (*hos doxei*)en to be of such an iron nature that he would stand fast in justice...

“How would you react?” the interlocutor is asked, and the thought experiment is under way.

Staying in ancient Greece, we have The Ship of Theseus thought experiment and its leading question about which of the two ships is identical with the original, the rebuilt ship or the one still sailing; its applications leading to the topic of personal identity, already present in early modern times with Locke and Hume.

On contemporary scene one of the most famous thought experiments is the Gettier one. Smith goes through various exercises, forms a belief, and the reader is confronted with the question “Does Smith know?” The thought experiment has revolutionized epistemology in the sixties.

Some years later Hilary Putnam (1973, 1975) has proposed a famous thought experiment: “Imagine a Twin Earth with water-like stuff of different chemical composition, say XYZ. Is Twin water water?” “No, since it is not H₂O.” The negative verdict about Twin Earth stuff is interpreted as revealing that, when it comes to our understanding of essential property and of stuff, composition is essential for stuff.

In philosophy of mind think of Searle’s Chinese room, with a human being, ignorant of Chinese language “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)” (Searle, 1999: 115). He gets incoming texts in Chinese, and, with the help of instructions, translates them, and then presents the answer in Chinese again. The whole, the-room-plus-person, would thus pass the Turing test: but does anybody (the person) or anything know Chinese here? No. But if our brain is like a computer, then it is like a Chinese room, and does not understand anything. But we do understand things, for instance, languages. So, our brain is not just a computer. “I demonstrated years ago with the so-called Chinese Room Argument that the implementation of the computer program is not by itself sufficient for consciousness or intentionality”, writes Searle (2010: 17).

On the side of ethics, we have to mention the Trolley problem, without retelling it, and on the side of political philosophy we have the Original Position as a “thought experiment for the purpose of public- and self-clarification”, as Rawls puts it in *Justice as Fairness: A Restatement* (2001: 17). The veil in the Original Position is “a guide to intuitions” (Rawls, 1971: 139). “The idea here is simply to make vivid to ourselves

the restrictions that it seems reasonable to impose on arguments for principles of justice, and therefore on these principles themselves” (Rawls, 2001: 16).

So much for examples. In the rest of the text we will briefly mention various topics connected to thought experiments. All of them are being discussed in the papers that follow, but not all to the same extent.

Firstly, what is a thought experiment? Einstein and Infeld wrote: “we recognized the importance of the idealized experiment created by thought” (1938: 226). The idealized experiment can never be actually performed, although it leads to a profound understanding of real experiments. They present the following characteristics of TEs:

- they are ‘fantastic’, they can neither be derived from experiment nor actually performed;
- they are consequently created ‘by thought’ and even by ‘speculative thinking’;
- they are however ‘consistent with observation’ and lead to ‘a profound understanding of real experiments’. (ibid.)

More recently, John Norton offered the following characterization of TEs:

“Thought experiments are arguments which: (i) posit hypothetical or counterfactual states of affairs, and (ii) invoke particulars irrelevant to the generality of the conclusion” (1991: 129).

Sören Häggqvist talks about “hypothetical cases intended to function as experiments, in the following sense: they aspire to test hypotheses or theories” (2009: 57).

A related question concerns the intellectual ‘relatives’ of TEs. J. R. Brown (1991, 1999, 2004) has proposed that some mathematical ways of reasoning are very close to TEs, people working in philosophy of literature (like Catherine Elgin, 2006, 2014) talk about some famous stories, like Orwell’s *1984*, in terms of TEs, and Descartes, in his discussion of *Meditations*, draws some serious analogies between thought experiments and religious meditations. Can this bewildering variety be somehow organized, and what is the wider genus to which TEs belong (the question is tackled briefly in the book)?

Once these questions are answered, one can tackle the most difficult part, the accounting and explanation of TEs and the presumed insights that they

enable us. J. R. Brown has, in *Laboratory of the Mind: Thought Experiments in the Natural Sciences* (1991), proposed a list of explananda, i.e. of challenges that TEs pose to the theoretician and philosopher of science. The first challenge is posed by the fact that in TE “there has been no new observational data” (Brown, 1991: 11). Other puzzles range from quickness to universality.

There are Platonist, empiricist argumentativist, Kantian, and naturalist-cognitivist proposals available for explanation. J. R. Brown has, on the contemporary scene, pioneered a Platonic proposal that imaginative intuitions are windows upon Platonic heaven. In the present volume, Majda Trobok comes closest to this view, in application to mathematical intuitions.

John Norton (2004) has famously been offering a more empiricist proposal: materials of a TE are empirical, and the method is just logical argumentation:

- (i) “My goal in this chapter is to state and defend an *account of thought experiments as ordinary argumentation* that is disguised in a vivid pictorial or narrative form.
- (ii) TEs invoke particulars irrelevant to the generality of the conclusion” (2004: 45).

Timothy Williamson agrees on the importance of logic, but disagrees on the role of particular cases: for him they are crucial in mobilizing our imagination, which then, together with abstract thinking, provide an answer.

The Kantian Buzzoni (2008) talks about “indeterminate multiplicity of categories individually constructed by the mind and tested *operationally* to solve specific problems” (2008: 15), ascribing to the categories a relative apriority.

On the naturalist-cognitivist side, the most popular proposal engages mental models (Nersessian 1992, 2007; Mišćević, 2007, 2017). Models are being discussed a lot in the literature. A recent account due to Chris Frith (2007) claims that in general, perception and action involve the generation of mental models. These models are continuously tested against reality and adapted using sensory signals and prediction errors. Although the basic mechanism enabling the construction and manipulation of the representations might be computational, the most important feature of representations and

operations having to do with mental models is precisely their concrete and quasi-spatial character. This has earned them the name of imaginative ‘mental models’. Mental models are typically involved in understanding stories or in ordinary planning of activities. When a reader encounters a description of a situation, she builds a model, a quasi-spatial ‘picture’ of it. As new details are supplied by the story-teller the model gets updated. The background conditions are dictated by the thought experimenter’s general knowledge about the world. Mišćević claims: “Much more important is the possibility of producing new data by manipulating the old ones. So, the manipulability of elements in the model and the mobilizing force of particular instances, more friendly to our ordinary reasoning process than the abstract general arguments, should account for the heuristic value of TEs. The theory of mental models makes the correct prediction that a visualizable TE will have more heuristic value than a corresponding propositional alternative, and explains in principle why. It is plausible that manipulating quasi-spatial models can in heuristic value come very close to actually seeing the movement of pieces manipulated” (Mišćević, 2007).

The next issue to be addressed in the explanation is the nature of abilities or capacities involved in thought-experimenting. A Platonist has to propose such special capacities; others might opt for a more naturalistic approach. One line of thought is to just claim that all our cognitive abilities contribute, and that in principle it is impossible to separate the contribution of the more specialized ones from the rest. This leads into ‘ordinarism’, the holistic view that would, on the justification side, treat the resulting insight as all *a posteriori*, or would question the very distinction between *a priori* and *a posteriori* justification (Williamson, 2013). On the opposite side is the view stressing particular competences, and thus allowing for some, say the mathematical ones, to yield *a priori* insights. M. Devitt (2013), himself more of an ordinarist, has ironically called it ‘Voice of Competence View’. This position is briefly defended by Mišćević in the present volume.

In recent times the debates just mentioned have been overshadowed by a new, big disagreement. A group of philosophers, most of them initially inspired by Stephen Stich, have decided to test the reliability of thought experiments, including the homogeneity of answers, dependence on irrelevant factors, and the like, in the way which is normally central in psychology. The new school has called itself experimental philosophy, X-phi for short. The experiments have produced some worrisome results that have inspired the creation of a more negative, critical stream within the school (E. Machery, R. Mallon, S. Nichols, S. Stich). Several lines of

defence of thought experiments against the accusations of negative experimental philosophers have been developed, a prominent one being the so-called expertize defence (T. Williamson, 2007; R. Sorensen, 1992). Miščević (2017) proposed an alternative line, building upon the distinction between stages of thought experiments and pointing to ways to avoid traps typical for each stage.

A final topic to be mentioned is the historical one. Thought experiments have resulted in traditions, some of them very long, for instance the Social Contract tradition (with the Original Position thought experiments starting as a sub-tradition within it), and the Cartesian sceptical tradition. Others, like the Twin Earth thought experiment or Gettier cases, have prompted at least half a century long trails. One cannot understand contemporary analytic philosophy without understanding these trails and traditions. It is an interesting question how the structure of a typical TE relates to the trail and what is the relation between its synchronic pattern and its diachronic posterity.

The book *Thought Experiments between Nature and Society*, a Festschrift for Nenad Miščević, is divided into three parts: Thought Experiment – General; Thought Experiments and Nature; and Thought Experiments and Society. Some of the texts are more explicitly about thought experiments, and some only glance at the basic problem, as they are more focused on other areas of Miščević’s philosophical contributions. Each text is accompanied by Miščević’s reply on the work of his colleagues.

The first part of the book (Thought Experiment – General) starts with **Timothy Williamson’s** text, **From Anti-Metaphysics to Metaphysics**, where he emphasizes that there is a more general moral about the famous ‘linguistic turn’, a phrase which has looked less and less appropriate as a description of mainstream analytic philosophy over recent decades. Nevertheless, although analytic philosophers are ceasing to regard their central questions as linguistic or even conceptual in any sense that would distinguish them from questions asked in other disciplines, the traces of the linguistic turn are not simply being erased. However, we learned from history that misconceptions in semantics often induce misconceptions in metaphysics. By causing fallacious metaphysical arguments to be treated as valid, coherent metaphysical views are incorrectly dismissed as confused or inconsistent. Therefore, a rich legacy of methodological sophistication with traces of the linguistic turn is useful in testing the soundness of arguments about non-linguistic matters, and analytic

philosophers regularly and justifiably draw on this work in both semantics and pragmatics.

Howard Robinson, in **Intuitions and Thought Experiments**, says that thought experiments are not at all in the speculative hunch category, like Dennett's imagined case. The issue with Mary is whether one can tell directly that the qualitative nature delivered in experience is a different property from, say, a neural firing structure, as represented in science. The issue relevant here is whether the kind of appeal to intuition involved in claiming to be able to identify the nature of a property by introspective awareness of its presence is, *prima facie*, question begging or illicit. For Robinson, it is clear that one cannot block the argument simply by saying it relies on an appeal to intuition, as if all such were illegitimate. In the classic 'mind' cases, the appeal is always to some version of the intuitively obvious, in a way that invites arguments challenging whether they are obvious in the way they seem.

Maja Malec and Olga Markič, in their text **Miščević on Intuitions and Thought Experiments**, begin with the assumption that, in the case of philosophical thought experiments concerning metaphysical modality, imagination has a more difficult task. We are often required to imagine situations that are very different from actuality, and what we end up with cannot be simply verified in actuality. Take, for example, Chalmers' thought experiment purporting to show that consciousness cannot be explained in physicalist's terms. We are invited to imagine our zombie twin, who is our physical duplicate, psychologically and functionally identical to us, behaving exactly like us, and being situated in the same environment like us, with the only difference being that our zombie twin lacks conscious experience. In other words, there is nothing it is like to be a zombie. If such a situation is imaginable and consequently possible, then physicalism is false. Obviously, if spatial perceptual constraints are not relevant here, what is? It is very simple to imagine my zombie twin, but is this imagined situation really possible? Clearly, I cannot actually create such a creature and thus confirm its possibility, but this does not mean that Chalmers' hypothesis is false. The matter stays undecided. Therefore, it is not enough that imagination is a good guide to actuality; it must be a good guide to possibility.

Nenad Smorkrović, in **Curiosity and the Argumentative Process**, begins with the idea that the issue concerning argumentative process has its source in early theoretical differentiation, going back to Plato and Aristotle, distinguishes between *rhetoric* and *dialectic*. Aristotle famously

made a difference between *rhetoric* and *dialectic*. He was the first who stressed the dialectical character of argumentation conceived as a critical discussion, in contrast to rhetoric, which is primarily motivated by finding the available means of persuasion. The intrinsic curiosity directs the dialogue towards the truth or towards the extension of knowledge. Being sincerely curious and acting in a joint effort to extend their knowledge concerning the truth of a proposition, the participants are motivated to do their best to achieve it. This means that in the AP, more than in any other format of reasoning, they are prone to de-bias their reasoning capacities.

Peter Gärdenfors - Semantic Transformations tackles the differences between metaphors and metonymies. Metaphors refer to mappings *between* domains and metonymies refer to meronomic and other relations *within* domains. For Gärdenfors, the role of metonymy is primarily referential. Metonymic concepts make it possible to conceptualize something by its relation to something else to which it is connected. By picking out one specific aspect of an entity, metonymy focuses attention on something salient to the situation, thereby helping one's understanding of it. The relevant parts and wholes in *pars pro toto* and *totum pro parte* metonymies need not be just spatial, but can be temporal as well. More abstract meronomic relations of a functional or causal nature also generate metonymies. He says that in the sentence "Proust is tough to read," the author stands for the book he wrote; while in "Napoleon attacked Russia," the highlight is Napoleon's function as the leader of the army. At the end, Gärdenfors' position is that by narrowing the use of the notion of a domain, one obtains a sharper division between metaphors and metonymies.

Danilo Šuster, in Lucky Math: Anti-luck Epistemology and Necessary Truth, has lumped several approaches in virtue epistemology together, since, according to his analysis, it is not clear how to augment the safety account with virtue-theoretic conditions. With the resources of virtue epistemology, why still stick with safety with all of its baggage (methods, ways of forming beliefs)? Why not virtues all the way? There are proposals for a virtue-oriented theory of luck on the market: one's belief is epistemically lucky when one fails to deserve credit for reaching the truth (Greco, 2003). Or, whether an agent is lucky depends on whether the agent is in control of her action, so one knows that p when one comes to have a true belief that p in a way that is sufficiently under one's control (Riggs, 2009). It seems that the approach of virtue epistemology will undermine the project that Mišćević started with: formulate an anti-luck condition in the frame of a modal minimal change theory of luck.

Guido Melchior - Epistemic Luck and Logical Necessities: Armchair Luck Revisited, starts with the claim that the particular problems that Mišćević and Pritchard encounter point towards a more general problem. In case of necessarily true propositions, we cannot consider possible worlds, which vary with respect to the truth value of the target proposition, but we can consider other variations, e.g. different cognitive apparatus (Mišćević) or different propositions (Pritchard). However, for such cases, we can modify the setting in a way that any possible world where the variation takes place is modally far off – and in this case, the necessary conditions for knowledge are fulfilled. Thus, modal variation does not offer a way out of the problem.

The last chapter in the first part is **Smiljana Gartner's Did a Particularist Kill the Thought Experiments?** A further (dis)advantage, especially of ETEs serving as intuition pumps, is that we can easily and quickly change them, and that some of them form very distant possible worlds. But where is then stability, one of Mišćević's desiderata? With changing ETEs, we are directly confronted with the complexity of our intuitions, situations, and ethical dilemmas. Let us take Foot's Tram Case. We can let the tram go its own course or we can redirect it. We can save five lives or one life. The intuition of the students says that we should divert the train and save five lives. But then we add the information that the one person is a baby and the five are 50-year-olds. Now they would choose the baby. Or we add the information that the baby has severe mental problems and that the five have small children. Now they would choose the five people again. By adding properties (perhaps to avoid the objection to under-description), the relevance of the features could change and, consequently, the conclusion could change as well, thus throwing stability out of the window.

The second part (Thought Experiments and Nature) opens with **Marian David's** paper, **Experimental Philosophy, Gettier-Cases, and Pragmatic Projection**. David presents an idea that there is a problem of principle in the approach of experimental philosophy. On the one hand, the experimenters want their 'intuition probes' to resemble the originals rather closely, otherwise they run the risk of missing their target. On the other hand, the originals are often perceived as rather odd, especially when taken out of the context of philosophy classrooms or textbooks, and they are usually rather sketchy. The situations they present are under-described in various ways (the philosophers who designed them put in only what they needed, which increases the risk that the subjects will fill in additional information when they encounter the cases in a survey setting –

consequentially they could arrive to unintended interpretations). However, if experimenters take measures to control against filling in, they run the risk that their ‘intuition probes’ will not probe untutored intuitions: the line between controlling for unintended interpretations and tutoring intuitions is fuzzy – and probing tutored intuitions is exactly what the experimenters do not want to do.

Peter Simons, in **Concepts in a World of Particulars**, starts his analysis with the following dilemma: If there is a discernible, identifiable abstractive basis to concepts, its *concreta* are either very complex, or very heterogeneous, or both. This is probably why it is so tempting to identify concepts with meanings. The problem is: although meanings themselves are less than straightforward compared with simple *abstracta* like numbers, masses, and shapes, their abstractive basis in the uses of words is fairly easy to identify. On the other side, concepts are involved with their sustaining acts, users, bits of language, activities, and compliants, and so on in a much messier and tangled way. This suggests that, despite the usefulness of talking about concepts, their status as putative *abstracta* is even shakier than that of ‘better behaved’ *abstracta*. Luckily for a nominalist, this causes no ontological anxiety: the sustaining individuals, within and outside the mind, are what matter and that do all the work.

Ihlan Inan, in **Is the Speed of Light Knowable A Priori?**, starts with the well-known idea that regarding any declarative sentence that expresses a proposition in a given context, one can know two things: the first is whether the sentence refers to a fact, and the other is the fact to which the sentence refers to (assuming that it is true). This requires us to distinguish between *knowing that a proposition is true* on the one hand, and *knowing the fact that makes a proposition is true* on the other. Knowing that, a proposition could then appear in two different ways. In the one case, the subject knows both that the proposition is true, and the fact that makes it true. This is ostensible knowledge. In the other case, the subject knows that the proposition is true, but does not know the fact that makes it true. This is inostensible knowledge. Illustrating this with a simple example: suppose a friend of yours says that his lover’s eyes are his favourite colour. If you take his word for it, you now know a proposition expressed by the sentence (in the appropriate context): “His lover’s eyes are his favourite colour.” Even though you know that the proposition is true, you may still not know what makes it true, that is, you may not know the fact to which the sentence refers. This would happen if you do not know your friend’s lover, if you do not know what is your friend’s favourite colour, or both. The sentence would then be inostensible for you, not because you

do not know whether it expresses a truth – you already know that it does – but rather because you do not know the fact to which it refers.

Andrej Ule, in **Mental Models in Scientific Work**, explains us that mental modelling in science seems to be an integral part of public practices of the usage of non-linguistic or para-linguistic representations of empirical structures in finding their full theoretical explanation. It is only secondarily a part of individual mental practices. The latter may be the result of internalization of the public practices by well-trained scientists. However, he believes that the same holds true for many types of intuitive insight. We have to ask ourselves in what case, in what circumstances, do we say: “Now I see what you mean.”, “Now I see how I can do this or that.”, etc.; “when *has* a formula, a graph, or a mental picture occurred to me?” Sometimes, when we hit upon the seemingly indubitable statements about intuitive reasoning, we actually stumble across indubitable, but *public criteria* for meaningful use of statements on our basic intuitions, and not on some *a priori* true propositions. Therefore, some innate (cognitive) competences and abilities may help in this development, but it will be a problem for further research if they are their neural or genetic foundations.

Ferenc Huoranski, in **Natural Kinds and Conceptual Truth**, presents, at the very beginning of the paper, a fictional example, which is, in fact, a real life example about cases when the content of a natural kind term is determined by its microphysical constitution that exists in abundance. Even if it is true of the concept of water or gold that their content was originally determined by water’s and gold’s sensible qualities and that we discovered their constitution only later, this is pretty much a contingent fact *not* about gold or water, but about how we actually introduce a term into our language. A procedure in the opposite direction is equally possible. Some chemical elements’ and compounds’ names have been first introduced by identifying their constitution, and hence it seems an *a priori* matter that their concept refers to some stuff with that type of constitution. Consequently, by what properties we fix a natural kind term’s reference – or how we provide cognitive content – should be irrelevant regarding the modal status of propositions involving the terms. For any moderately realist account of the nature of modal truth, it must be unacceptable that a truth’s modal status is merely determined by such historical accidents.

Majda Trobok, in **Grasping the Basic Arithmetical Concepts - The Role of Imaginative Intuitions**, tries to explain and criticize Miščević’s position about how to come to the notion of natural number. For him,

Hume's principle – taken informally – does play a crucial role. However, any attempt to show that such a principle is, in fact, intuitively clear, or that the grasping process of it is based on quasi-perceptual (or imaginative) intuition, is, Trobok suggests, highly controversial, since there is no practical-empirical or historical-mathematical support based on evidence for it. She says that even though Miščević does not ignore the mathematical practice, he does not look at the historical-mathematical side of it, deciding to concentrate rather on what developmental psychology has to say on how we grasp the notion of (natural) number. She tries to go one step further and claims that we do not have sufficient practical-empirical or historical-mathematical evidence in favour of the view that grasping such concepts is based on imaginative intuitions, neither for the concept of equinumerosity (Frege), nor for that of (natural) numbers.

Andraž Stožer and Janez Bregant, in **The Colour Dilemma: A Subjectivist Answer**, at the beginning delineate the neurophysiological mechanisms underlying the visual experience of colour, and then with the help of concrete results of practical neurophysiological experiments show that the physicalist and dispositionalist theories of colour, in the light of new, neurophysiological objections, do not present credible views on the nature of colour. These findings demonstrate that a direct electrical activation of neurons that usually respond to blue-purple stimuli suffices to evoke a perception of blue-purple colour by the subject. Therefore, answering the question of what colour really is, physicalism and dispositionalism should be, in the light of neurophysiological findings that speak in favour of colour not being out there in the world, replaced by subjectivism as the theory of colour that addresses the problem of colour realism based on scientific grounds.

Matjaž Potrč - Dasein in a Vat says that it would be wrong to start from a sceptical conclusion about the external world. There is a tendency (e.g. Searle) to conclude that the assertion of mine and my BIV counterpart's experiential world identity leads to external world scepticism, and accordingly to the denial of external world existence. For Potrč, this is a wrong conclusion. Notice that the author, hopefully and by presupposition not a BIV, has an experiential world. If he perceives a cat in his experiential world, his thought about the cat will be true. Whereas his BIV counterpart's thought will be false, for the counterpart has no real world similar to the one that the author is operating in. Identity of our experiential world (based upon sharing of the zero point essential indexical) does not imply identity of our external worlds. It is simple: author's BIV, in counterpart to himself, has no such external world at all,

he just shares the experiential world together with the zero point perspective with the author.

The third part of the book (Thought Experiments and Society) opens with **Pierre Jacob's** paper, **Knowing One's Own Mind**. His starting point is that even if preverbal infants can spontaneously represent the contents of others' false beliefs, most 3-year-olds may fail explicit change-of-location false-belief tasks and point to the toy's actual location for at least three complementary reasons. First of all, they might fail to compute the relevance of the agent's false belief for answering the experimenter's prediction question. Secondly, even if they do compute its relevance, they might lack the executive resources required for inhibiting their own *knowledge* about the marble's *actual* location from intruding into their answer to the experimenter's prediction question. Finally, they might misinterpret the experimenter's prediction question as a normative question about where the mistaken agent ought to look for her toy. Note further that in order to predict where the mistaken agent will look for her toy, it is sufficient to know where she last placed it. Whatever happens to the toy afterwards is irrelevant. So in the classical explicit task, children are provided, by the experimenter, with much information about the whereabouts of the agent's toy that is irrelevant to predicting the agent's action. In a nutshell, having the ability to track the contents of others' false beliefs is not sufficient for success at elicited-response false belief tasks. Therefore, failure at these tasks by 3-year-olds is not a demonstration that they lack a theory-of-mind. Nor does it show that 3-year-olds cannot think that they and others think.

Friderik Klampfer, in **The False Promise of Thought-Experimentation in Moral and Political Philosophy**, shows that one important reason for distrusting intuitions generated by thought experiments is their malleability, instability, and vulnerability to manipulation. Our intuitions are easily swayed one way or the other by simple rephrasing of the story, a change in the order of presentation, emotional and social priming, or simply by tampering with our physiological needs. What psychologists tell us about the mechanisms that produce them and what we know influences them doesn't exactly build confidence. Intuitions are quick, snap, unreflective, spontaneous, almost automatic judgments. If they are preceded by any reasoning at all, it must be subconscious. They rely, for their formation, on similar cognitive shortcuts, heuristics, which people use in their judgments in other domains (availability, representativeness). They are subject to the framing effects ('lives not saved' vs. 'lives lost') and moralizing spill-over effects and in many cases they are shaped by

mood, affection, emotion, fatigue, affected by the level of abstraction and sensitive to the order of presentation. It is important to stress that despite their contingent origin, they are mostly dogmatic, i.e. resistant to contrary evidence. When our intuitive judgments are challenged or questioned, we are seldom able to provide good reasons or compelling evidence in their support (or if we are, the reasons we adduce are often not those that were operative in the production of our judgment). Furthermore, Klampfer states that: we fail to see any need for that and, consequently, don't consider this to be a problem (what is called 'moral dumbfounding').

Miomir Matulović, in **Miščević, Mental Models, and Thought Experiments in Political Philosophy**, shows in his analysis that Miščević's proposal of the desiderata for good or successful political thought experiments sets very high standards that political thought experiments have to satisfy. It seems that in this way Miščević is trying to answer the critics of the method of thought experiments in political philosophy. But he is aware that these standards (fulfilment of the desiderata) are never fully satisfied by the actual, even most famous, thought experiments. Did the critic thus win? Or is Miščević thinking of a so far unformulated alternative? For the time being, it seems that Miščević is confronted with a dilemma: either raise the standards, in order to satisfy the critics of thought experiments, or admit that standards are never satisfied by the extant proposals. On the first horn, the critics might be satisfied with the demands but will point to the fact that these are never fulfilled, and on the second horn Miščević has to agree with the critics of thought experiments.

Boran Berčić, in **Are Nations Social Constructs? Nenad on Nations**, starts with the proposal that the idea that nations are social constructs can be spelled out in the following way: Individuals *a, b, c, etc.* form a nation *A*, iff they believe that they form a nation *A*. The relevant question here is what does it mean that *they* believe? What is a collective belief? It seems that it is not sufficient that each individual believes that he or she belongs to the nation. One has to believe that others have the same belief as well. Therefore, it might be said that individuals *a, b, c, etc.* believe that they form a nation iff: (1) each one of them believes that they form a nation, and (2) each one of them believes that every one of them believes that they form a nation. The next question that arises is what does it mean that they *believe*? One might say that it is sufficient that they behave *as if* they believe (1) and (2), for in that case we might ascribe corresponding beliefs to them. However, in this case, large herds or flocks of animals could satisfy the requirement. Therefore, it seems that one has to demand that

they must have some degree of rational reflection about their belonging to a nation. Berčić says that they must have a disposition to give an affirmative answer if asked whether (1) and (2).

Rudi Kotnik, in **Thought Experiments in Teaching: TE as a Suppositional Real Story**, thinks that there is a contrast between two ‘sorts’ of TEs: a) TE as an imagined story; and b) TE as a suppositional real story. They are both TEs, but with one single, tiny, but crucial difference: in the case of b) students believe (or are supposed to believe) that the story is real. It seems that we are justified to assume that their intuitions will be real. If (when) they are real, we can expect ‘real’ motivations, engagements, passions – and corresponding positions that they defend (in accordance to their intuitions). He explains that we do not ask students: “Imagine that refugees are on their way to our city, neighbourhood, and even to our school, etc.,” but we say: “They are already coming etc.” For sure, it is a fake information, perhaps morally questionable, but still an experiment and an interesting learning experience. If we would ask them: “Imagine that they come here etc.,” they would, probably, respond in a politically correct manner, and the discussion about this issue and these kinds of issues of politics and morality would be quite (or perhaps even significantly) different. Therefore, we did not say: “Imagine that they come,” but we said: “They are already coming.” And the responses are expectable: “Yes, we understand the situation, but we cannot accept them.”; “Why don’t they send them to some other place?” Their real intuitions, revealed through the exercises, become a possibility to reflect upon them in contrast to the usual political correctness.

Boris Vezjak, in **The Ring of Gyges and the Philosophical Imagination**, asks this question: Is Plato’s story of Gyges only a paradigmatic example of philosophical imagination and nothing else? What does it mean to say that? Imagining things is a part of philosophical activity and in itself has always been a tool for philosophizing. The dilemma is whether such imagination is only supplementary, or does it have a more fundamental part to play in constituting and developing philosophical arguments. Was Plato’s idea about picturing the consequences of invisibility really meant to be something stronger than mythically inspired exercise in fictionalising? TEs have to be much more than simple exercises in fictional activity. It appears the interpretation about ethical TEs went beyond the bounds of Plato’s possibility to conceptualize TE.

Reference list

- Brown, J. R. (1991). *Laboratory of the Mind: Thought Experiments in the Natural Sciences*. London: Routledge.
- (1999). *Philosophy of Mathematics: An Introduction to the World of Proofs and Pictures*. London: Routledge.
- (2004). “Peeking into Plato’s Heaven”. *Philosophy of Science*, 71, pp. 1126–1138.
- Buzzoni, M. (2008). *Thought Experiments in Natural Science: An Operational and Reflexive Transcendental Conception*. Würzburg: Königshausen & Neumann.
- Devitt, M. (2013). “Linguistic Intuitions are not ‘the Voice of Competence’”. In Haug, M. C. (Ed.), *Philosophical Methodology: The Armchair or the Laboratory?* London: Routledge. p. 268.
- Einstein, A., Infeld L. (1938). *The Evolution of Physics*. Originally from Cambridge University Press. Reprinted by Simon and Schuster. New York, 1966.
- Elgin, C. (2006). “The Laboratory of the Mind”. In Abel, G. (Ed.), *Kreativität: XX Deutscher Kongress für Philosophie, Kolloquiumsbeiträge*. Hamburg: Felix Meiner Verlag, pp. 772–784.
- (2014). “Fiction as Thought Experiment”. *Perspectives on Science*, 22, pp. 221–241.
- Frith, C. (2007). *Making up the Mind: How the Brain Creates Our Mental World*. London: Wiley-Blackwell Publishing.
- Greco, J. (2003). “Knowledge as Credit for True Belief”. In DePaul, M., Zagzebski, L. (Eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press.
- Häggqvist, S. (2009). “A Model for Thought Experiments”. *Canadian Journal of Philosophy*, 39 (1), pp. 55–76.
- Miščević, N. (2007). “Modelling Intuitions and Thought Experiments”. *Croatian Journal of Philosophy*, 7 (2), pp. 181–214.
- (2017). “Accounting for Thought Experiments – 25 Years Later”. This book.
- Nersessian, N. J. (1992). “In the Theoretician’s Laboratory: Thought Experimenting as Mental Modelling” In Okruhlik, K., Hull, D., Forbes, M. (Eds.), *PSA 1992: Proceedings of the Biennial Meeting of the Philosophy of Science Association, Vol. 2*. Chicago: The University of Chicago Press, pp. 291–301.
- (2007). “Thought Experiments as Mental Modelling: Empiricism without Logic”. *Croatian Journal of Philosophy*, 7, pp. 125–161.

- Norton, J. (1991). "Thought Experiments in Einstein's Work". In Horowitz, T. Massey, G. (Eds.), *Thought Experiments in Science and Philosophy*. Lanham: Rowman & Littlefield, pp. 129–148.
- . (2004). "On Thought Experiments: Is There More to the Argument?". *Philosophy of Science*, 71, pp. 1139–1151.
- Plato. (1991). *The Republic*. Translated by B. Jowett, B. Vintage Classics edition.
- Putnam, H. (1973). "Meaning and Reference". *Journal of Philosophy*, 70 (19), pp. 699–711.
- . (1975). "The Meaning of 'Meaning'". In *Mind, Language, and Reality: Philosophical Papers* (Volume 2). Cambridge: Cambridge University Press, pp. 215–271.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge, Massachusetts: Harvard University Press.
- . (2001). *Justice as Fairness: A Restatement*, Kelly, E. (Ed.). Cambridge, Massachusetts: Harvard University Press.
- Riggs, W. (2009). "Luck, Knowledge, and Control". In Haddock, M., Pritchard, D. (Eds.), *Epistemic Value*. Oxford: Oxford University Press.
- Searle, J. (1999). "The Chinese Room". In Wilson, R. A., Keil, F. (Eds.), *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, Massachusetts: MIT Press, pp. 115–116.
- . (2010). "Why Dualism (and Materialism) Fail to Account for Consciousness". In Lee, R. E. (Ed.), *Questioning Nineteenth Century Assumptions about Knowledge, III: Dualism*. New York: SUNY Press.
- Sorensen, R. A. (1992). *Thought Experiments*. Oxford: Oxford University Press.
- Williamson, T. (2007). *The Philosophy of Philosophy*. London: Blackwell.
- . (2013). "How Deep is the Distinction between A Priori and A Posteriori Knowledge?". In Casullo A., Thurow, J. C. (Eds.), *The A Priori in Philosophy*. Oxford: Oxford University Press, pp. 291–312.

CONTRIBUTORS

❖ **Boran Berčić, University of Rijeka**

Boran Berčić earned his doctorate from the University of Zagreb. He is now Professor of Philosophy at the University of Rijeka, Croatia, where he teaches the courses Introduction to Philosophy, Metaphysics, and several electives. He has written a number of articles and five books: *Znanost i istina (Science and Truth)*, HKD, Rijeka, 1995; *Realizam, relativizam, tolerancija (Realism, Relativism, Tolerance)*, HKD, Rijeka, 1995; *Filozofija Bečkog kruga (Philosophy of the Vienna Circle)*, KruZak, Zagreb, 2002; *Filozofija I (Philosophy I)*, Ibis grafika, Zagreb, 2012; *Filozofija II (Philosophy I)*, Ibis grafika, Zagreb, 2012.

❖ **Bojan Borstner, University of Maribor**

Bojan Borstner is a Slovenian philosopher. He received his doctorate at the University of Ljubljana, and is now Professor of Philosophy and Head of the Department of Philosophy at the Faculty of Arts, University of Maribor, Slovenia. His main research areas are ontology, philosophy of science, and the philosophy of nursing. He has published two books, *Problems of Realism* and *Virtual Teacher* (with B. Aberšek and J. Bregant), more than a hundred articles and shorter pieces, and edited three books – *Handbook of Phenomenology and Cognitive Science*, *Consciousness at the Crossroads of Philosophy and Cognitive Science*, and *Historicizing Religion: Critical Approaches to Contemporary Concerns*.

❖ **Janez Bregant, University of Maribor**

Janez Bregant is a Slovenian philosopher, with a Ph.D. from the University of Maribor, Slovenia, where he is now Associate Professor. His areas of interest are philosophy of mind and philosophy of artificial intelligence; his areas of competence are philosophy of art and critical thinking.

❖ **Marian David, University of Graz**

Marian David obtained his Ph.D. in 1990 from the University of Arizona, Tucson. After 20 years as Professor of Philosophy at the University of Notre Dame, Indiana, he now holds the Chair for Theoretical Philosophy at Karl-Franzens-University of Graz, Austria. He is co-editor of the journal