Thought Experiments: State of the art

Michael T. Stuart, Yiftach Fehige and James Robert Brown

Introduction

Hans-Christian Ørsted introduced the term "thought experiment" to philosophy in 1811, although it did not get much philosophical attention until the 1980s. This is therefore a relatively young topic of discussion (in philosophical timescales), still concerned with the most fundamental questions of what thought experiments are, what they do, and how they do it. This introduction will trace some of the work that has been done already. The chapters that follow will take us further.

An enumeration of some classic examples will display how diverse the category of thought experiments can be.[1] In philosophy, there is Searle's Chinese room, Putnam's twin Earth (and brain in a vat), Nozick's experience machine, Rawls' original position, Jackson/Dennett's colour scientist, Thomson's violinist, Chalmers's zombies, Wittgenstein's beetle, Plato's cave (and ring of Gyges), Quine's *gavagai*, Davidson's Swampman, Poincaré's diskworld, Foot's trolley problem and many more.[2] These thought experiments in large part define the history of philosophy and are woven into its pedagogy. Indeed, there is a textbook (now in its fifth edition) that aims to introduce students to philosophy entirely through thought experiments (Schick and Vaughn 2012). And their power to engage minds often extends beyond philosophy classrooms: there are several collections of thought experiments aimed at wider public audiences (e.g., Tittle 2004; Cohen 2004), and they appear frequently in educational contexts online.

Thought experiments are equally common in science. Some famous examples include Lucretius' throwing a spear at the edge of the universe, Maxwell's demon, Einstein's elevator (and train), Schrödinger's cat, Newton's bucket (and cannonball), Heisenberg's microscope, Galileo's falling bodies (and pendulums, inclined planes and ship), the Prisoner's Dilemma, and Stevin's chain draped over a prism. They are found also in pure and applied mathematics, where they play important roles from geometry to infinity.[3]

While philosophical and scientific thought experiments have stolen most of the scholarly spotlight, the scope of thought experiments grows substantially when we recognize the many works of art that we can fruitfully characterize as thought experiments. These might include paintings such as Jackson Pollock's "Number One"; novels such as Mark Twain's *Huckleberry*

Finn, Harper Lee's *To Kill a Mockingbird*, Sophocles' *Oedipus Rex*, Charles Dickens' *A Tale of Two Cities*, Vladimir Nabokov's *Lolita*, George Eliot's *Middlemarch*, William Shakespeare's *Henry V*, *King Lear* and *Hamlet*, George Orwell's *Animal Farm*, Harriet Beecher Stowe's *Uncle Tom's Cabin* and E. M. Forster's *Howards End*; and films including *The Matrix* and *2001: A Space Odyssey* (see Carroll 2002; Wartenberg 2007; Camp 2009; Davies 2012, this volume; Elgin 1993, 2014). It might also be useful to discuss the creation and appreciation of art as a kind of thought experiment.

Besides humanistic, scientific and artistic pursuits, thought experiments are deeply important for ordinary life (Nersessian 2007). On an inclusive conception of "thought experiment," simply planning out a busy day might require several; figuring out how best to get from one place to another, deciding what to eat, etc. Thought experiments and repeated visualizations may also be used in therapy, for example, to overcome phobias (Gendler 2004).

Thought experiments thus form an extremely diverse set of mental activities. Since there is no conventional definition, we should resolve to begin our inquiry with as broad a notion as possible. That is, we should be aware of and minimize our pre-theoretical definitional criteria until we have exposed ourselves to the full diversity of cognitive activities that might warrant the label "thought experiment." Only when it comes time to develop historical, sociological and philosophical accounts of thought experiments should we narrow our focus to the subsets and properties of thought experiments that interest us.

We therefore invite the reader to treat this book as a companion, not in the sense of an expert guide who can navigate unfamiliar terrain to a predetermined destination, but in the sense of an enthusiastic fellow explorer. Our first step (to be taken in Part I) exposes us to many of the different types and aims of thought experiments.

Thought experiments in their historical contexts

One important approach to thought experiments is historical, the value of which far outstrips any use philosophers make of it. We want to know what role Ibn Sīnā's (Avicenna's) thought experiments played in his epistemology and how they were received by his contemporaries independently of what such an inquiry might tell us about thought experiments in general. Focusing on thought experiments merely provides historians with another useful lens through which to study thinkers and periods that are already worth studying. And in Part I we look through this lens.

Still, it would be wrong to underplay the close connection between the history and philosophy of thought experiments. Generally speaking, historians will not be able to catalogue and analyze historical thought experiments without making philosophical assumptions about the nature of inference, evidence, imagination and their mutual relations. Conversely, philosophers have no hope of producing a general account of thought experiments without the examples and insights of historians. In particular, historical studies such as Gellard (2011), Ierodiakonou (2005, 2011), Knuuttila and Kukkonen (2011), Kühne (2005), Lautner (2011) and Palmerino (2011) have raised important philosophical issues that might otherwise have gone unnoticed. For instance, the history of thought experiments shows that many have acquired lives of their own, despite Ian Hacking's (1992) claim to the contrary. This is an important datum for developing ontologies and taxonomies of thought experiments (see Bokulich and Frappier, this volume). The interdependence of the history and philosophy of thought experiments is clearly reflected in the chapters of Part I, which may also serve as an extended introduction to thought experiments and (with Part II) a source of evidence against which to test the philosophical claims made in Parts III and IV.

Given the ubiquity of thought experiments, we could not attempt to provide comprehensive historical coverage. In the literature to date, there has been a focus on thought experiments in Western traditions, mostly following Galileo and the birth of what is called modern science. James McAllister (this volume) argues that this focus is appropriate since thought experiments are a tool introduced by Galileo with a specific sort of evidential significance linked to the Western notion of scientific experiment. For the reasons given above, we have sought chapters that go beyond this conception. Plato might be using his imagination in a way that is not the same as Galileo's, but similar enough that we might learn something new about thought experiments or thought experiment–like inferences by considering them. So in Part I, we tentatively expand our notion of thought experiment, while still covering the cases that are most discussed in the literature. We hope this provides a balance between historical scholarship on well-known thought experiments and those that are less well-known. In the first category we have thought experiments in the presocratics, Plato, Galileo, Newton and Leibniz. In the second we have Aristotle, Islamic/Arabic thinkers, Kant, Wittgenstein and Hegel. This selection should provide readers who are new to the topic of thought experiments the historical background necessary to appreciate the rest of the companion, while readers already comfortable with the history of thought experiments will find new scholarship that opens original directions for research. In terms of expanding the literature, we have only gone a small fraction of the way. Thinkers from feminist philosophy, African, Asian, Latin American traditions and others should also be represented in the discussion, as well as disciplines that have not traditionally received the attention of philosophers, including architecture, engineering, law and advertising, which are all likely to have a great deal to contribute.

Before moving on, we would like to provide a brief summary of the chapters of Part I. Katerina Ierodiakonou in her chapter on the ancients argues that thought experiments were used for three

main purposes: to support, to attack and to induce suspension of judgment on philosophical claims. She considers three illustrative episodes: Archytas of Tarentum's throwing a spear at the edge of the universe, the myth of Gyges in Plato's *Republic*, and Sextus Empiricus' "partless places." Rather than discussing these episodes as thought experiments, Ierodiakonou helpfully refers to them as *paradeigmata*, or "examples." To illustrate how such examples were marshalled, Ierodiakonou considers a few more cases, including the ship of Theseus, Chrysippus' Dion and Theon, and the skepticism-inducing arguments of the stoics and Skeptics. In his chapter on Plato, Alexander Becker draws attention to Plato's many literary devices and styles in order to ask how Plato's fictional creations compare to thought experiments and how we are supposed to learn from them. Becker focuses on the *Republic*, which includes not only what we might classify as thought experiments but also a discussion of what we can learn from fiction. He identifies three main fictional styles in Plato – myth, simile and dialogue – and relates these to the myth of Gyges, the myth of Er, the allegory of the cave, the myth told to the citizens of the model city, the construction of the model city, and the *Republic* itself as a dialogue.

Klaus Corcilius begins his chapter on Aristotle as Ierodiakonou and Becker do: by admitting there is no simple way to read into Aristotle our notion of thought experiment. And this is no accident, argues Corcilius, since Aristotle regarded plain-sight observation as superior to the "artificial extraction of hidden facts" (this volume, 58). However, "there can be no doubt," Corcilius argues, that Aristotle did use what *we* call thought experiments. Their principle function seems to be to "compensate for a lack of data in epistemically difficult terrain" (60). To illustrate what this means and how Aristotle achieves it, Corcilius considers examples from both Aristotle's esoteric writings (dense philosophical notes) and his exoteric (literary) writings. The latter include a cave allegory similar to Plato's and a frightening analogy of the soul's relation to the body in terms of a living person chained to a dead one. The former include "the stripping argument," which analyzes substance itself, imagining if there were a second sun, a consideration of "if all things were colours" (or if everything were only one colour), why flesh is not the organ of touch, and finally what moves the universe and how.

Jon McGinnis discusses medieval Arabic/Islamic thought experiments by discussing fictions, idealizations and "ingenious machines," primarily in the work of Ibn Sīnā, but also in Ibn al-Haytham, Abū Ḥamid al-Ghazālī and others. These philosophical tools were put to a variety of uses, including supporting modal claims and illustrating ideas. McGinnis shows that the use of hypothetical conditionals, a tradition going back to Aristotle, was beginning to be recognized as philosophically problematic, especially in connection to its dependence on the imagination. McGinnis follows this with a fascinating discussion of an internal sensory faculty introduced by Ibn Sīnā to avoid objections about hypothetical counterfactuals called *wahm*, which is a faculty separate from imagination and intellect that enables us to perceive non-sensible features or intentions in sensible particular things (as when a sheep perceives the ferocity of a wolf, or I "see how you feel"). This faculty was crucial for explaining our ability to grasp mathematical objects, work with idealizations and employ thought experiments. McGinnis goes through Ibn Sīnā's

famous "flying man" thought experiment and many others to show the creativity and subtlety with which thought experiments were used and criticized during this period.

In striking contrast to the usual way Galileo is presented, Paolo Palmieri rejects the distinction between real and thought experiments for Galileo. Instead, Galileo's "experiential engagements" are recast as "projective participations." One way to put the point is that recent recreations of Galileo's experiments reveal that his experiments, if material, could not have shown what Galileo concluded from them. Therefore, the experiments must have been conceived and completed always (at least partially) projectively in thought. Palmieri's second main point is that Galileo was not trying to dissolve paradoxes with his experiments. He was celebrating them. To show this, he displays some of Galileo's most beautiful struggles with falling (and floating bodies).

Richard T. W. Arthur fruitfully groups together some deservedly famous thought experiments of Isaac Newton and Gottfried Leibniz. These include Newton's bucket and the Leibniz shifts, wherein God creates the universe in a different place, in the same place but rotated, or at a different time. The unifying theme of all these thought experiments is the use of indistinguishability to show something about the nature of motion and matter. Arthur discusses the context and motivations of these thought experiments in relation to the work of other philosophers like Descartes, Hobbes and Locke, and shows how historically important theses such as Leibniz's identity of indiscernibles were originally grounded in indistinguishability thought experiments about matter and motion. The discussion branches out to other indistinguishability thought experiments concerning God's abilities, the nature of perception and the nature of mind, one of which anticipates Searle's Chinese room thought experiment.

In Chapter 7, Kenneth R. Westphal charts the development of transcendental thought experiments found in Kant, Hegel, Wittgenstein and others. Westphal prefaces this with a historical-philosophical discussion of the nature of philosophy and its limits, especially with respect to the possibility of global perceptual skepticism and massive reference failure. Through this preface, Westphal breathes life and immediacy into his discussion of these thought experiments which are transcendental in the sense that they try to grasp some of our key cognitive capacities and their limits. Examples include Kant's thought experiments that ask us to conceive of a complete absence of space or time; Kant's thought experiment in which there is no regularity in the sensory manifold; and Wittgenstein's thought experiments in *On Certainty* which ask us to violate in imagination some of the most fundamental regularities we know to obtain, such as conservation of mass. Despite their apparently external subject matter, these are thought experiments about what we can and cannot think. Our failures or difficulties become crucial data for the construction of philosophical theory.

The glimpse provided by Part I into historical uses and examples of thought experiments naturally leads to certain philosophical and sociological questions such as: What are the features

common to (sets of) thought experiments? Do different communities draw different divisions between thought experiments, fictions, models and arguments? How contextual are the success criteria for thought experiments, and what causes a community to change them? Answering these questions requires a somewhat broader perspective, and a natural thing to do is to look at *fields* of inquiry rather than periods and thinkers. For example, what are the sorts of thought experiments used in politics, economics, theology, ethics, physics, biology and mathematics? What unites them, how are they used and how have they changed? This is the goal of Part II, which collects domain-specific observations and presents accounts of the characteristics and practices of thought experimentation in the fields just mentioned. This part of the collection is organized like the first, to include chapters that summarize and expand on previous work (e.g., in physics, mathematics and ethics) with some relatively new areas ripe for cultivation (economics, biology, theology and politics). As with the previous section, further expansion is desirable, for example, to thought experiments in metaphysics, chemistry, geology, sociology, engineering, law, architecture, etc. A brief summary of these chapters follows.

Thought experiments in their disciplinary contexts

Nenad Miščević (co-originator of the mental models account discussed by Nancy Nersessian in Chapter 17) discusses thought experiments in political philosophy. These typically present some social arrangement, ask us to judge it, and then use that judgement as evidence for a political or philosophical claim. From the model city (also discussed by Becker, this volume), we learn about justice itself. From behind Rawls' veil of ignorance, we judge different arrangements of "the basic structure" of society and reason about which is best. Similar thought experiments are found in Rousseau, Dworkin, Cohen and Kukathas, and many more are collected in Tetlock and Belkin's (1996) *Counterfactual Thought Experiments in World Politics*, which helped to spur recent literature on the topic. Miščević considers the characteristics and stages of such thought experiments, analyzes historical and contemporary examples, and argues even for the indispensability of such thought experiments in political thought by canvassing the possible sources of evidence available to political philosophy. Finally, Miščević outlines some general desiderata for successful political thought experiments and asks what this discussion can tell us about ongoing philosophical issues, such as the status of ideal theories in political reasoning.

Some authors claim or imply that economics proceeds primarily by thought experiment. Taking up a contrary position, Margaret Schabas argues that not only are thought experiments rare in economics, but when we do find something that looks like one, it usually isn't. Her position relies on a conception of thought experiments according to which at least some experimental manipulation is necessary (which in turn relies on a particular notion of experiment). According to this conception, a mathematical model that operates by derivation, no matter how hypothetical

its content is, cannot be a thought experiment. Schabas also rejects the view that modelling (and mental modelling) is experimental. Schabas does identify a few genuine thought experiments in economics, such as Hume's, in which everyone in a society wakes up with an extra five units of currency in their pocket. And she is happy to admit that economics is increasingly experimental. However, the "thought" and "experiment" portions of economics are not very often combined in the way that we see, for example, in physics. Some of these claims are taken up by Julian Reiss (this volume), who attempts to explain why there are less thought experiments in economics than elsewhere.

Thought experiments have long been part of the method of theology, but the features of such thought experiments and what their involvement says about theology are open questions. Yiftach Fehige considers six ways thought experiments and theology have or might come together in philosophical discourse. First, there are some who dismiss the use of thought experiments in ethics for theological reasons. Second, their appearance in theological contexts has been taken as evidence against the reliability of thought experiments in general. Third, the presence of thought experiments in both science and theology has been used to argue that there are comparable standards of rationality in these two disciplines. Fourth, theological thought experiments have an intriguing literary nature that might be helpful in advancing the investigation of thought experiments through aesthetics (as in, for example, Davies, this volume; Elgin 2014; and Meynell, this volume). Fifth, the claim that philosophy would be impossible (or at least severely impoverished) without thought experiments also appears to be true of philosophical theology. Finally, what is called revealed theology has its own thought experiments, which are fundamentally different from those we find elsewhere.

This brings us to ethical thought experiments, which are among the most well-known in any discipline. Like philosophers of politics (if Miščević is correct), ethicists seem required to appeal to thought experiments because normative claims are not amenable to the same sources of evidence as less normative disciplines. In his chapter, Georg Brun provides a general characterization of ethical thought experiments, including the infamous trolley problem, Thomson's violinist, Singer's drowning child, Shue's ticking time bomb diffused by torturing a terrorist, and Rawls' original position. Brun produces a typology and considers objections (including moral objections) to the use of thought experiments in ethics.

Kent A. Peacock turns us to physics, which is the only discipline that could rival philosophy in terms of the quantity and historical significance of thought experiments. Though it is an open matter whether thought experiments in ethics, epistemology or metaphysics achieve their goals, the question in physics is not if, but how. Peacock is more liberal in his characterization of thought experiments than Schabas, including even textbook problems that deal with simplified scenarios like blocks sliding down inclined planes. Peacock notes that there are thought experiments that illustrate or draw attention to certain ideas, but he chooses to focus on those that have left indelible marks on physical theory and practice well beyond the intentions of their

creators. These include Maxwell's demon, the Einstein-Podolsky-Rosen thought experiment, Einstein's riding a light wave, Einstein's elevator (and its progenitor – an observer falling off a roof), Einstein's "hole" argument, Plank's oscillators, Einstein's mirror, Heisenberg's microscope, Schrödinger's cat and even quantum computation. In all of these cases, theoretical solutions to the problems raised by thought experiments, and problems with the theoretical solutions suggested by those thought experiments, make up important parts of the history of physics. Peacock concludes with a discussion of the future of thought experimentation in physics.

Unlike physics, biology is not brimming with famous thought experiments. Guillaume Schlaepfer and Marcel Weber begin their chapter by considering why this might be. For instance, the theoretical portion of biology seems to play a "more marginal" role in biology compared to the role played by theoretical physics in physics. And biology, unlike physics, has more direct experimental access to its systems. This does not mean that there are no interesting thought experiments in biology, however. For example, Darwin's work contains many. Schlaepfer and Weber consider Jim Lennox's (1991) account of these thought experiments and raise some interesting questions. For example, if thought experiments should involve the imagination, *how much* imagination do we require for something to count as a thought experiment? This is especially poignant in the case of Darwin, whose thought experiments are characteristically full of references to empirical evidence. Does this make them less imaginative (and therefore less thought experimental)? At this point the chapter turns to the work of mathematically focused population geneticists. In this literature we find R. A. Fisher's sexually reproducing community with three (or more) sexes, and other candidates for biological thought experiments.

 Finally, Irina Starikova and Marcus Giaquinto tackle the question of thought experiments in mathematics, where, as they point out, it is not at all strange to see our stock of knowledge increase from thought alone. A more interesting issue, they claim, is the use of sensory imagination in mathematical thought experiments. They provide several examples of such thought experiments from knot theory, graph theory and geometric group theory. Many of these are justified because the visualizable manipulations of mathematical objects (like knots) are tied directly to formal manipulations. "For foundational purposes there needs to be some way of fixing the subject matter in mathematical terms, so that the correctness of basic assumptions and methods can be proven. But once that job has been done, we may proceed without adverting to our foundational definitions" (this volume, 262). In their exploration of graph theory, they find visual imaginings that proceed "in a truly experimental way" (272). They go on:

> The utility of visual imagination depends on confining our efforts to images and image transformations which are simple enough for us to manipulate reliably in imagination. But the variety of images and image transformations that we can handle reliably suffices to make visual imagination a potent instrument of mental experimentation in mathematics. (272)

In sum, there are thought experiments in mathematics that involve "active use of visual imagination" and go "beyond the application of mathematically prescribed rules, as a way of answering questions or overcoming obstacles" (275). Intriguingly, the epistemic support provided by these thought experiments is deemed *empirical*. This might seem counterintuitive, but "empirical evidence has a much larger role in the epistemology of actual mathematical belief acquisition than is often thought" (276).

With some understanding of the history and disciplinary context of thought experiments, we turn to Part III, which presents and reappraises the main existing philosophical accounts of thought experiments. Our contributors develop, defend and criticize epistemological accounts that attempt to explain when and how a thought experiment succeeds. There is a deflationary empiricist account that treats thought experiments as a mere subset of arguments (Chapter 15), a rationalist account that treats thought experiments as possible stimulants of rational insight (Chapter 16), a naturalist account that portrays thought experiments as the manipulation of what cognitive scientists call "mental models" (Chapter 17), a transcendental account that portrays thought experiments both as continuous with real experiments and also as conditions for the possibility of real experiments (Chapter 18), and a phenomenological account that takes as fundamental the first-person experience of performing a thought experiment (Chapter 19).

To introduce this section, we think it might be helpful to have a quick look at the history of these accounts.

Some history of the philosophy of thought experiments

Unfortunately, there are only a few resources concerning the history of the philosophy of thought experiments (e.g., Kühne 2005; Moue, Masavetas, and Karayianni 2006; Fehige and Stuart 2014), and those sources focus mostly on philosophy of science. This can perhaps be forgiven, since the earliest notion of a *Tankeexperiment* relied for its meaning on the scientific experiment. This is equally true for Ørsted as it was for Georg Christoph Lichtenberg, Novalis, and Immanuel Kant, who were writing just before Ørsted using similar notions (see Schildknecht 1990; Daiber 2001; Fehige and Stuart 2014). The philosophers who followed Ørsted, including Ernst Mach, Alexius Meinong, Pierre Duhem, Thomas Kuhn, and Imre Lakatos, continued to focus on the experimental nature of thought experiments.

Mainstream philosophers picked up the moniker in the mid-1980s. Jonathan Dancy considered the use of thought experiments in ethics, and Daniel Dennett considered them in epistemology and metaphysics under the name of "intuition pumps." Dancy and Dennett weren't convinced by thought experiments that purported to establish certain philosophical conclusions. Skepticism

intensified with the introduction of experimental philosophy.[4] But whether the discussion takes place in metaphilosophy or philosophy of science, the main epistemological worry is the same: can these thought experiments really be doing what they seem to be doing? This question takes its present form in the chapters of Part III through a dialogue that we think started with Kuhn.

Unlike Mach and those before him, Kuhn wrote about thought experiments mainly as a tool for motivating or justifying scientific revolutions. For Kuhn,

> A crisis induced by the failure of expectation and followed by revolution is at the heart of the thought- experimental situations we have been examining. Conversely, thought experiment is one of the essential analytic tools which are deployed during crisis and which then help to promote basic conceptual reform. (Kuhn 1977, 263)

Kuhn cites Einstein's train, Heisenberg's microscope, and several fragments from Galileo as examples of thought experiments that play this role in theory change. He calls these "an important class of thought experiments" (260–61), and he concludes that "from thought experiments most people learn about their concepts and the world together" (253).

For Kuhn, revolutionary thought experiments were not used to generate new facts, but to ease scientists through the arational period of crisis that exists between scientific paradigms, back to the rational progress of what Kuhn calls "normal science." In a period of crisis, we must weigh the competing claims, methods and potentials of rival paradigms. And it seems that thought experiments can and have helped us partially to transcend the confines of our paradigms, which is necessary if we are to be convinced of a new world-view. Kuhn argues that by changing and exploring our world-views by means of thought experiments, we learn indirectly about the world as well.

Kuhn's answer to the question of how thought experiments fuel scientific progress did not win widespread acceptance, although there is still some sympathy (for example, Sorensen 1992; Gendler 1998; Van Dyck 2003). What is important in the present context is Kuhn's idea that thought experiments play a justificatory role in science, and especially in scientific revolutions.

 This idea was central for those who organized the first conference on thought experiments in 1986, and it has been a point of contention ever since. The proceedings of that conference were published in Horowitz and Massey (1991), and on the first page of their introduction the editors point out that what is at stake is a "paradox" inspired by Kuhn's paper. The paradox of thought experiments consists in the "puzzling fact that thought experiments often have novel empirical import even though they are conducted entirely inside one's head."[5]

Kuhn's "puzzling fact" only became a "paradox" via the ensuing debate between James Robert Brown and John D. Norton, which presented two clear but conflicting solutions: one with "epistemic magic," and one without. Brown and Norton assumed with Kuhn that thought

experiments could play a justificatory role in scientific revolutions, and they both took the scientific record as their main source of information. However, they disagreed about what thought experiments were and what they could do. Brown presented a Platonic theory of thought experiments and Norton developed an empiricist account that characterized thought experiments as arguments. Brown claimed that thought experiments could provide direct access to truths about laws of nature, which is something Norton derided as magical. Each was attempting to resolve Kuhn's puzzle.

Brown began with Galileo's falling bodies thought experiment (see Palmieri, this volume) in which

> we have a transition from one theory to another which is quite remarkable. There has been *no* new empirical evidence. The old theory was rationally believed before the thought experiment, but was shown to be absurd by it. The thought experiment established rational belief in a new theory. (Brown 1986, 10)

It does this independently of experience because "there has been no new observational data" (11), the conclusion does not follow deductively from the premises (11–2), and because "it is not a case of seeing old empirical data in a new way" (which is, Brown writes: "essentially Kuhn's thesis" 1986, 11). So Brown accepted that at least some thought experiments perform the role that Kuhn envisaged, "a crucial role in paradigm change" (2). They play this role by providing reasons to reject one theory and adopt another, and those reasons are not strictly logical, nor do they rely solely on previous sense-experience.

Norton also focused on revolutionary thought experiments, and he agreed with Brown concerning the problem: "Thought experiments in physics provide or purport to provide us information about the physical world. Since they are *thought* experiments rather than *physical* experiments, this information does not come from the reporting of new empirical data." But he drew a very different conclusion:

> There is only one non-controversial source from which this information can come: it is elicited from information we already have by an identifiable argument … The alternative to this view is to suppose that thought experiments provide some new and even mysterious route to knowledge of the physical world. (1991, 129)

Norton thus presented Kuhn's puzzle in the form of a dichotomy, bringing us to the paradox: if thought experiments provide new information about the physical world, yet do not require new information about the physical world, either the new information is a rearrangement of old data, or else it comes from rational insight.

A great deal of papers and books published on thought experiments since 1991 in philosophy of science have mentioned or focused on this paradox.[6] Most of them present Kuhn's problem in

slightly different terms, or call it by a different name. Given that the paradox is rarely presented in the same words, some writers have begun to question whether there is actually more than one paradox.

For instance, Horowitz and Massey characterized the paradox in terms of the "fact that thought experiments often have novel empirical import even though they are conducted entirely inside one's head." Depending on how we interpret "novel," "empirical import," and "entirely inside one's head," we get different versions of the paradox. There are many ways for something to be novel, for example, by being surprising, non-derivable, or by presenting us with a previously non-existing belief, experience, ability, pattern, property, or relation. Likewise, there are many ways of having empirical import, for example, by prompting a change in our existing set of empirical beliefs, knowledge, evidence, understanding, information, etc. Finally, to be "inside one's head" might require that the *elements* manipulated in a thought experiment be in the head, that the thought experimental *process* be in the head, or that the *evidence* that justifies the output of a thought experiment be in the head (for an extended discussion of these and other options, see Stuart 2015, ch. 1 §§3–5).

For Brown and Norton, empirical import meant empirical knowledge. Brown and Norton disagreed, however, concerning the relevant sense of novelty and also concerning the sources of evidence in the heads of thought experimenters.

Before considering where other philosophers in the literature stand with respect to novelty, empirical import and independence of experience, we'd like to summarize the two chapters in this volume dedicated to the views of Brown and Norton. In her chapter, Elke Brendel helpfully distinguishes five main sub-claims of Norton's position. Using this framework, she explains the various attacks and defenses that have been made, and others that could be made, against each of them. She identifies one of these claims as impossible to reject without begging the question against Norton, and presents several powerful arguments against three others. One of the claims, the *Reconstruction thesis*, is argued to be defensible. Brendel concludes by showing that if we reject what she regards as the weaker of Norton's claims, this will not commit us to Brown's position.[7]

In his chapter, Thomas Grundmann focuses on the two main arguments Brown gives for his position: 1) an inference to the best explanation, according to which Platonism is the best explanation for the practice of scientific thought experimentation, and 2) an argument from analogy that ties Platonism in mathematics to Platonism in science. Both of Brown's arguments rely on a particular account of the laws of nature, namely, the Dretske-Tooley-Armstrong account. According to this account, laws of nature are relations between universals. If it is correct, perhaps Brown is right that empiricism cannot explain our mental access to such universals and their relations. And since the rational access to the mathematical realm is *also* a matter of access to universals, we gain support for the analogy between mathematical and

scientific thought experiments. Grundmann forcefully attacks both of the two main arguments, in addition to the account of laws of nature that Brown relies on. He then considers the ways in which Brown's account could be revised, and closes by asking whether and how a Platonic account of thought experiments could deal with philosophical thought experiments.

Faced with the extremes of rationalism and empiricism, many who contributed to the debate after 1991 sought a middle ground. Sören Häggqvist (1996, 2009) agreed with Norton that insofar as thought experiments are to play an evidential role in science or philosophy, they must participate in arguments. That is, thought experiments are used to contest or bolster theoretical claims by providing (usually modal) evidence that counts for or against a claim (see Williamson 2005, 2016 for a similar view). For Häggqvist, a thought experiment thus plays a justificatory role in the same way a real experiment does: by contradicting or supporting a claim made by the theory through participation in an argument. This makes (thought) experiments *parts of* arguments. However, Häggqvist denies Norton's claim that the performance of a thought experiment just is the performance of an argument, because an experiment cannot be formally valid or invalid.[8] Häggqvist's insight is taken up by Tim De Mey (2003), who argues that we should investigate the epistemic impact of the thought experiment's conclusion in one way, and how it produces that conclusion in another.

In other words, there is the conclusion of the thought experiment, which is a product of psychological mechanisms including imagination, memory and intuition. And there is the use of that conclusion in an argument for or against the truth of a claim, which might be justified in the standard ways identified by logicians. It is therefore important to be clear which of these two very different characterizations of the paradox we are addressing in any epistemological account of thought experiments: the reliability of the processes that bring us to the conclusion of the thought experiment, or the reliability of the inferences we use to make that conclusion bear on a theory or theoretical claim.

There are also naturalists in the debate who claim that we can and should use science to discover how thought experiments work. This idea was already present in Mach (1905) and Wolfgang Yourgrau (1962, 1967). But Roy Sorensen was the first to give an in-depth naturalist account of thought experiments (1992).

Like Häggqvist and Williamson, Sorensen portrayed thought experiments as a type of modal reasoning. And like Mach (1905), he placed thought experiments on a continuum with real experiments. Along with several others (including Richard Arthur, Alisa Bokulich, Tamar Gendler, Kuhn and Lichtenberg), Sorensen argued that thought experiments mostly eliminate irrationalities in our thought. And again following Mach, he claimed that thought experiments function by drawing upon the stores of empirical knowledge that we accumulate in our lifetimes, combined with the innate ideas and structures that have been programmed into our minds by

evolution. This is what makes Sorensen a naturalist in particular: his use of evolutionary psychology to justify the reliability of thought experiments.

Sorensen agreed with Norton that thought experiments "repackage" old information to make it "more informative" (1992, 4). In this sense, he adopted Norton's notion of novelty. But his interpretation of empirical import is quite different. Instead of empirical knowledge, he took the goals of thought experiments to include the creation and stabilization of phenomena, atheoretical exploration, and the definition of concepts. Achieving these goals does not usually amount to creating new empirical knowledge in the sense of providing new justified beliefs. One important point to learn from Sorensen about the paradox is that empirical import can mean much more than empirical knowledge. Catharine Elgin (1993, 2000, 2002, 2004, 2006 and 2014) has been focusing on thought experiments with respect to understanding instead of knowledge for some time. Mike Stuart (2016a, 2017, this volume) develops this idea as well.

Since Sorensen, more naturalists have emerged, many of whom rely to some extent on the argument from evolution. An important subset of these characterize thought experiments as "mental models," a technical term from cognitive science first applied to thought experiments independently and simultaneously by Nenad Miščević (1992, 2004 and 2007) and Nancy Nersessian (1992a, 1992b, 2007, 2008, this volume), and later by Tamar Gendler (2004) and others. Miščević and Nersessian agreed that "thought experiment descriptions" guide us in the creation of dynamic mental models which "mobilize" cognitive skills that render the outcome of the model manipulation epistemically efficacious. For example, we can make mental models that help us to produce reliable guesses about how and when water will spill out of variously shaped containers whose openings are tilted increasingly towards the ground. Such models draw on our everyday experience with containers and liquids and pouring. It is easy to see why many problems are easier and faster to solve when represented in a mental model as opposed to verbally or formally: we get to use the same abilities and knowledge we use in everyday real- life situations (like moving in a gravitational field), and this is (supposed to be) very different from working out a solution using logical inferences, as Norton would have it.

In more detail, Nersessian argues that the narrative presentation of a thought experiment triggers the creation of a mental model, which is "a structural, behavioral, or functional analog to a real-world phenomenon" (this volume, 311). The mental model is analogous in the sense that it preserves something of the constraints operating on the objects and events of the imagined phenomenon (311). Such mental models are often visual or tactile in nature, are manipulated in real time, draw on embodied wisdom (1992, 294), and embed a specific and personal point of view into the model (1992, 295). With respect to the paradox, Nersessian made a telling remark early on:

> The constructed situation, itself, is apprehended as pertinent to the real world in several ways. It can reveal something in our experience that we did not see the import of before

> … It can generate new data from the limiting case … [and] it can make us see the empirical consequences of something in our existing conceptions. (1992, 296)

In other words, like Sorensen, Nersessian recognized many of the different ways empirical import could be interpreted. Instead of producing new knowledge, a thought experiment can highlight old data that did not initially seem important, it can separate phenomena that seem necessarily connected, it can generate new data from limiting cases, and it can clarify the consequences of previous conceptual commitments.

In her chapter in this volume, Nancy J. Nersessian begins with a presentation of the varied advancements in the literature on mental models from psychology, cognitive science, philosophy and neuroimaging studies. After all, if we're going to take seriously the suggestion that thought experiments are mental models in the scientist's sense, we must understand what the scientists are saying. And once we do, there are many remarkable details concerning how we reason using models that can inform the discussion, or pose new problems. Nersessian closes by considering how the account relates to or subsumes other accounts already mentioned in this introduction.

There are many important insights to gain from the naturalists. In terms of novelty, Gendler, Nersessian and Miščević all agree with Norton that existing knowledge can be manipulated, transformed or rearranged in a thought experiment to draw our attention to something we didn't notice before. However, these philosophers also identify additional sources of novelty, including truly novel mental presentations, which could take the form of new experiences, concepts or beliefs. But unlike Brown, these would not be interactions with real, mind- independent abstract entities. Gendler, Nersessian and Miščević also agree that rearrangement can help us to "possess" our beliefs by giving us the power to act on them, as when a thought experiment helps us to overcome a fear of flying that we know is irrational. Statistical knowledge that flying is safe is not enough to prevent fear in some people, yet thought experiments in the form of repeated positive visualizations can help agents to make their statistical knowledge about the safety of airline travel useful (Gendler 2004, 1160). This sense of novelty concerns our abilities, and the relationships between our beliefs.

One final point concerns the distinction between propositional and non-propositional cognitive processes. Norton claims that "The actual conduct of a thought experiment consists of the execution of an argument, although this may not be obvious, since the argument may appear only in abbreviated form and with suppressed premises" (2004b, 50). Norton could be correct that the best way to explain the epistemological efficacy of thought experiments is to reconstruct them as arguments, which are manipulations of propositions, without being correct about the additional claim that the actual conduct of a thought experiment is the execution of an argument. Nevertheless, for Norton to be correct about this additional claim, we need to know if non-propositional cognitive processes are necessary for "the actual conduct" of a thought experiment. Mental modellers typically claim that such processes are necessary, or they reject the distinction

altogether. (On this issue see Peijnenburg and Atkinson 2003, 121–5, and Goffi and Roux, this volume). Others have sided with Norton, arguing that we need only make reference to manipulations of propositions for a descriptively accurate epistemological account of thought experiments (e.g., Salis and Frigg forthcoming). This issue remains open.

In addition to empiricism, rationalism and naturalism, there is a Kantian account of thought experiments that has been developed in detail by Marco Buzzoni (2008, 2010, 2011, 2013a, 2013b, 2015, 2016 and this volume). According to Buzzoni, thought experiments should be considered from two philosophical levels of analysis. According to the first, thought experiments are a condition for the possibility of real experiments. That is, without thought experiments, we could not see objects and events as answers to questions that we put to nature. In order to question nature, we have to imagine different ways the world could be (2008, 116–20). Without our ability to distance ourselves cognitively from the actual situations in which we find ourselves, we would not be able to think of how things *might* be, and so we could not generate questions or hypotheses about what *is*. Therefore, not until we are in a position to entertain different hypotheses about what the world is really like (as opposed to how it merely appears) does experimental practice become possible. In this sense, thought experiments and real experiments are *not* the same; thought experiments are more fundamental. According to the second level of analysis, however, we note that thought experiments are given their content by experience: they put questions to nature, anticipate answers using induction and deduction, and must ultimately be justified by appeal to experience. In this sense, thought experiments are on a par with real experiments. To put it in a slogan: real experiments without thought experiments are blind, and thought experiments without real experiments are empty.

One insight that results from this view is that we must be careful to differentiate between thought experiments as cognitive actions necessary for the scientific enterprise in general, and on the other hand as mental variations of variables that anticipate how nature will answer our questions. Failing to do this, Buzzoni argues, will force us to dismiss what is common to thought and laboratory experiments (as Norton must), or to misidentify thought experiments as epistemologically indistinguishable from laboratory experiments (as Brown's account seems to allow).[9] Buzzoni's account can accommodate both views. In his chapter in this collection, Buzzoni updates his account, first by responding to objections made by Yiftach Fehige (2012 and 2013), and then by mapping the terrain of possible Kantian approaches to thought experiments based on different interpretations of the Kantian a priori.

Another option is to produce a phenomenological account of thought experiments (see e.g., Mohanty 1991; Kujundzic 1995; Froese and Gallagher 2010; Fehige and Wiltsche 2013; Wiltsche 2013, this volume; Hopp 2014). Phenomenology takes as fundamental the "phenomena," that is, how things appear in conscious experience. One major insight of phenomenology is that what appears to us always outstrips our given sensory experience of it. Perceiving an object is not just receiving certain patterns of light on the retina; it requires

recognizing the object as something with various properties and a certain history, which would appear different from different perspectives or under different conditions, and so on. Given this, our experience of any object (material or otherwise) consists also of our expectations, background knowledge, and abilities. Edmund Husserl called this expanded notion of experience when applied to an object, "horizontal givenness." Harald A. Wiltsche (this volume) applies this phenomenological framework to thought experiments, first by providing an excellent introduction to phenomenology in terms that anyone can understand, and second by pointing out that it is not only material objects that have horizontal givenness, but imaginings as well. Drawing on imagined cases from *The Simpsons* as well as special relativity, Wiltsche presents a general but careful consideration of how the action of performing a thought experiment can shed light on the different ways that objects and events appear in the imagination, how our background knowledge figures into the action of thought experimentation, and how we can learn from the process.

Besides rationalism, empiricism, naturalism, Kantianism and phenomenology, there are still other positions that are possible to take, for example, pluralism, contextualism and skepticism.[10]

 And there can be combinations of existing approaches. Walter Hopp is a phenomenologist who treats thought experiments as fictions that give us rational access to universals (Hopp 2014), and one could be a Kantian who, when it comes to assessing individual thought experiments, is an empiricist. But we want to stress that despite the form of the above narrative, the philosophy of thought experiments is not a game of philosophical bingo in which we wait for all the philosophical -isms to be filled in and then go home. Thought experiments represent a rare four-way intersection of history, philosophy, cognitive science, and social science, and the opportunity for mutual information is precious. Here, knowledge from each of these disciplines (and others) can come together to enhance our understanding about the powers and limits of the mind. Each of the philosophical -isms can therefore be expected to inform but also *be informed* by historical discoveries, and cognitive and social scientific research.

The chapters in Part III represent much of the current state of affairs in the philosophy of thought experiments. Naturally, however, these chapters do not (aim to) exhaust all the possibilities. Addressing the questions they raise and also those that they don't is the purpose of Part IV.

Future directions

One lacuna in the above discussion is the lack of engagement between mainstream philosophy and philosophy of science. The first several chapters of Part IV help to address this. In the first, Stephen Stich and Kevin Tobia discuss the role of intuition in philosophical thought experiments

and the impact of experimental philosophy. They begin by differentiating intuitions stimulated by thought experiments that are used as evidence for claims about "in-the- head mental entities" (like the content or extension of concepts, or the warrant for implicit or tacit theories) versus "outside-the-head non-psychological entities" (like universals, modal or moral truths, or natural kinds). Their primary question echoes the paradox/puzzle mentioned above: how can intuitions about philosophical thought experiments be evidence for claims about extra-mental entities? While we lack a good story for why such intuitions might be thought to be reliable sources of evidence for extra-mental entities, experimental philosophy provides a set of powerful reasons to be skeptical about those intuitions. Stich and Tobia introduce and summarize what they take to be the major relevant insights of experimental philosophy, and then turn to some criticisms. They consider and reject the "expertise" defense, according to which the intuitions of professional philosophers are less vulnerable to irrelevant factors, and another defense that restricts what we mean by "intuition" in a way that would enable us to avoid the charges of experimental philosophy. This latter defense is one taken up by Kirk Ludwig, author of the second chapter of Part IV.

Kirk Ludwig asks what experimental philosophy would have to be like if it were to justify the skeptical conclusions of Stich and Tobia and others. Ludwig breaks experimental philosophy down into several sub-kinds, presents the main contemporary criticisms of experimental philosophy, and argues that experimental philosophy as it stands cannot cast serious doubt on the use of thought experiments in philosophy. It can display variation in intuitions, but this does not demonstrate the further claim that philosophical intuitions are mistaken, or that folk intuitions are correct. According to Ludwig, these further claims cannot be addressed using survey-based methods. Of course, this does not mean that experimental philosophy is not useful for philosophy; on the contrary, it can alert us to important cognitive biases and defects, it can help us design better thought experiments, and be better teachers to our students. Ludwig closes with a five-point reply to the criticisms of Stich and Tobia.

Sticking with thought experiments in philosophy, Sören Häggqvist and Daniel Cohnitz discuss a specific and very common *kind* of philosophical thought experiment, namely, the counterexample. For these thought experiments to succeed, they typically require the evaluation of at least one modal claim, usually a counterfactual, which raises the following question (another echo of the paradox): what reason(s) do we have to think that thought experiments can be reliable guides to modal claims?[11] To address the problem, several formal reconstructions of the use of thought experiments as modal defeaters are considered, as well as the possibility of skepticism raised by experimental philosophy, and insights from the literature on the inference from conceivability to possibility (e.g., Gendler and Hawthorne 2002). Like Stich and Tobia, Häggqvist and Cohnitz vote for pluralism: different subfields of philosophy might require different things from thought experimental intuitions, and different epistemological accounts might be necessary. They finish by urging a closer connection between the discussion of thought experiments in metaphilosophy and philosophy of science. In both subfields, they argue, the

dialectical context of a thought experiment appears necessary for a full epistemological account. This is an idea also championed by Goffi and Roux in Chapter 24.

In the next chapter, James W. McAllister considers cross-cultural and historical comparisons of thought experiments to "problematize the category of thought experiment and its application in different historical and cultural contexts" (425). He does this by asking *where* different communities draw the line between thought experiments and other kinds of reasoning, *when* they use thought experiments, and under *which* conditions thought experiments are taken to have evidential significance. These issues are closely connected to the metaphilosophical issues of the previous three chapters, to the metaphysical concerns of the final chapter, but also to the entire first two parts of the collection, in which we find displayed exactly the sorts of differences McAllister examines. Based on these observations, McAllister argues against using the term "thought experiment" for any inference occurring before Galileo. In so doing, McAllister takes the contextual nature of thought experimental justification even further than Cohnitz and Häggqvist and Stich and Tobia. He employs a fully historicized view of evidential significance, according to which something counts as evidence *only* in a given socio-historical context. Because of this, ancient Greeks, for example, would not have accepted thought experiments as a method with evidential significance since they would not have accepted experiment in general as such a method. McAllister considers the use of thought experiments in the discipline of history as well as in the philosophical traditions of India and China to provide additional reasons against imperialistically labelling various imaginative activities as thought experiments.

 Taking seriously the contextual and dialectical nature of thought experiments suggested by the previous chapters, Jean-Yves Goffi and Sophie Roux present a dialectical account of the logic of thought experiments. They argue that all thought experiments require the entertainment of a scenario, the mental manipulation of which requires non-propositional knowledge (knowledge-how, tacit knowledge, etc.). Meanwhile, for thought experiments to produce propositional knowledge, they must also trade in propositions. However, no propositional reconstruction could ever be completely filled in, and this renders thought experiments irreducibly "opaque." According to them, a more complete account must admit the importance of propositional reasoning while accounting for this irreducible opacity. They argue that dialectical accounts of reasoning were created to deal with exactly this sort of opacity, since in the dialectical exchanges of natural conversation we operate without agreed upon rules or definite background assumptions. They examine three different ways dialectic argumentation theory has been developed in the twentieth century, and argue in favour of Nicholas Rescher's account (1977), which they apply to thought experiments.

When we want to circumscribe the epistemological power of thought experiments, we often focus on successful instances. But an equally fruitful way of proceeding is to concentrate on those that fail. One very helpful example is given by John D. Norton in Chapter 25, "The Worst Thought Experiment." This is a thought experiment proposed by Leo Szilard in 1929, which was

presented as a development of Maxwell's demon. Through a characteristically clear historical account, Norton identifies two important features of our interaction with thought experiments. First, we allow "extensive latitude" to the creator of a thought experiment in introducing idealizations, which are necessary to focus the narrative. Second, we typically presume with the creator of the thought experiment that the case considered is a typical one, from which an inductive generalization can follow. Norton explains the failure of Szilard's thought experiment in terms of its (mis)use of both of these features.

Following Norton's discussion of idealization in Szilard's thought experiment, Julian Reiss analyzes the role of idealizations in thought experiments in general. First, he distinguishes several kinds of idealizations, and argues that only one is relevant to this discussion: what Ernan McMullin called "Galilean idealizations." Of the subtypes of Galilean idealization, Reiss concentrates on causal idealizations, or in other words, those that "isolate a single causal line." Reiss asks how and under which conditions such idealizations can be justified. He identifies two conditions, one epistemic and another empirical. To satisfy the epistemic condition, there must be some way to know what would happen in the idealized single causal line. One way is to look at an asymptotic series of material idealizations and infer to the extreme case, like experimenting with increasingly smooth objects to infer to the behaviour of frictionless ones. To satisfy the empirical condition, the idealization must be informative concerning the behaviour of the same causal line when in the non-idealized context of the real world. Reiss argues that these two conditions explain why we see more thought experiments in some domains than others.

Reliance on idealizations is one of many features thought experiments share with computer simulations. Their similarity has inspired strong claims, such as the prediction that in some domains, computer simulations will replace thought experiments altogether (Chandrasekharan, Nersessian, and Subramanian 2013). In Chapter 27, Johannes Lenhard reviews and tackles various comparisons between thought experiments and computer simulations. Despite the similarities, Lenhard argues that the production of intuitions is importantly different from the automated iterations of formal algorithms, and that this has epistemological ramifications. Most notably, computer simulations can be opaque in the sense that they are not completely accessible to epistemic assessment, and this is because there is a type of iteration they employ that would be impossible in a thought experiment. This sort of iteration is displayed in Schelling's model of social segregation, the Ising model, complex systems models like dynamic meteorological models, and models that approximate the Schrödinger equation for systems of molecules too complicated to solve directly. Because of the differences in iteration-type, different epistemological accounts will be needed to explain the different strengths and weaknesses of computer simulations and thought experiments.

Lenhard concludes by discussing the common historical roots of thought experiments and computer simulations. For one thing, both emerge from a natural need and ability to explore hypothetical possibilities. Whatever we call this ability in the silicone domain, in humans we

tend to call it imagination. In Letitia Meynell's "Images and Imagination in Thought Experiments" (Chapter 28), we get an in-depth look at the imaginative core of human thought experiments. Part of Meynell's project is to bring together insights from the literature on imagination in philosophy of mind and aesthetics. Through a discussion of how we interact imaginatively with images, Meynell considers the strengths and weaknesses of images in the context of thought experimentation. For instance, Einstein's train is accompanied by an image in the original text that might not be necessary to justify the conclusion of the thought experiment, but which serves as an extra "prop" (in the sense of Walton 1990) that directs us to imagine scenarios in certain ways, reassures novices that they are imagining the right things, and foregrounds certain aspects of the imaginary scenarios as well as certain skills and subsets of background knowledge. And because there aren't any norms dictating how we must approach a given image, for example, from right to left or top to bottom, we are encouraged to work through it ourselves, which compliments the experimental nature of thought experiments. Meynell closes with some epistemological considerations about how to justify uses of the imagination in thought experiments.

There are other ways to bring aesthetics to bear on the topic of thought experiments. For example, Ichikawa and Jarvis (2009) argue that it is not just background assumptions that we rely on in a thought experiment, but our ability to interact with *stories*. Perhaps the human brain has evolved some reliable way of forming modal inferences from imagining what would be true in a fictional world, and thought experiments take advantage of this. Such an account would work nicely with the mental models account, since fictionalism about scientific models has also been a popular topic in philosophy of science for some time, and if thought experiments are used as models are, then insights from that literature can be brought to bear.

David Davies presents just such a combination. An on-going question in philosophy of art concerns how we learn from fiction. One might think that all literary and cinematic fiction could do in order to increase our stock of knowledge would be to *suggest* hypotheses for empirical study. That is, we might think that works of artistic fiction are (at best) confined to the context of discovery. On the other hand, we don't (always) make this same claim for the thought experiments found in science and philosophy. A number of issues are pertinent to a defense of the possibility of learning through fiction by comparison with thought experiments. For one, in science we make explicit claims about the world, and there are (at least conventional) epistemic standards by which we can judge whether those claims are true or warranted. This does not seem to be the case in fiction, as fictions typically do not make explicit claims about the world and there are typically no standards by which to judge the epistemic value of their content. These are the "no-argument" and "no-evidence" problems, respectively. That is, fictions make no arguments that the world is a certain way, and they provide no evidence independent of what the reader herself brings to the fiction. After arguing against several influential accounts of learning through fiction, Davies presents his account of what it means to be a fictional narrative, and argues that thought experiments in science and philosophy satisfy it. He then provides a helpful

taxonomy of epistemological views on the issue of learning from fiction ranging from extreme deflationism to extreme inflationism, and selects "moderate inflationism" as the best option. This view is meant to cohere with the mental models account of thought experiments, according to which thought experiments mobilize unarticulated cognitive resources to learn something new about the world from fiction. Davies argues that this survives the "no-evidence" objection because through such mobilization, a fiction can make us aware of patterns that "underlie the complexity of prior and present actual experience" (this volume, 521). And it survives the "no-argument" objection because we test the fiction's claims *through* our engagement with it, so we need not appeal to anything external to the fiction. Several interesting questions remain, however. For instance, how do we apply this moderate inflationist account to thought experiments in film, which are not "verbally mediated" as most thought experiments are?

Like Meynell and Davies, Michael T. Stuart also seeks to connect the discussion of thought experiments to insights in another philosophical subdomain. In this case, we are directed to the literature on understanding. Stuart argues that the epistemological challenge of thought experiments is equally interesting when phrased in terms of understanding as when phrased in terms of knowledge. He proposes three ways that thought experiments can increase understanding of x: by explaining x, making x more meaningful, and making x more fruitful. He argues that these different ways of enhancing understanding can yield the different types of understanding identified in the epistemological literature. To address the question of how thought experiments produce understanding, Stuart draws on new work in epistemology and aesthetics by Elizabeth Camp and Alison Hills.

In the final chapter, Alisa Bokulich and Mélanie Frappier critically appraise various definitions of thought experiments to see if anything definite can be said. Bokulich and Frappier must deal with several related challenges: First, what makes two thought experiments similar enough to be counted as repetitions of the same experiment? Second, how do we accommodate the insight that *what counts as equal* will itself be relative to the purposes of classification? That is, for some purposes historical continuity is what matters, for others it is propositional content or structural form. Third, and closely related: do thought experiments change over time, and if so, how? To answer these questions, Bokulich and Frappier take up a pluralist stance. They argue that what counts as identical depends on what we think thought experiments are. If they are intuition pumps, identical thought experiments should produce the same intuitions. Given the variability of intuitions mentioned by Stich and Tobia and Cohnitz and Häggqvist, perfect identify between two thought experiments might be rare. If thought experiments are sets of premises and a conclusion, thought experiments cannot change or be modified (except in the order or presentation of the premises, or in the irrelevant but picturesque details of the narrative). If thought experiments are mental models, then the thought experiment outstrips its linguistic presentation. While there are worries about what might count as equal for the (essentially private) mental models, there might be enough to give thought experiments identities which can be modified over time. If thought experiments are props for the imagination (Salis and Frigg

forthcoming; Meynell 2014, this volume) then for each thought experiment there are two sets of fictional truths, one that is objective and socially shared, and one that is created in the subjective imagination of each participant using their own principles of generation (see Meynell, this volume). Since these sets differ, we can explain why two different conclusions can be drawn from the same thought experiment: the principles that generated the intersubjective sets of fictional truths were different, or improperly applied by one party, or the same but insufficiently clear. This will not describe all cases, however, since in many it is not just the principles of generation that are different or variously applied, but the narrative itself that is changed, sometimes drastically. Finally, if thought experiments are experiments, they will be replicated just like other kinds of experiments. Again, however, there are different aspects of experiments that can be replicated. This brings us again to the "dual- structure view" of thought experiments (originally inspired by Hans Radder's dual view of laboratory experiments (1996) which was developed by Häggqvist and De Mey. However Bokulich and Frappier point out that this account might still be insufficient to capture certain important cases.

Concluding thoughts

We hope Part IV of the collection furthers the existing debates about thought experiments, and creates conceptual space for new competing accounts and issues. As a whole, we hope the companion invites discussion, criticism and comparison in as many new directions as possible.

Before concluding, we want to emphasize again that the distinctions between the parts of this companion are not, and could not be, strict. The chapters in Parts I and II present new research, much of which pertains directly to the soundness of the arguments in Parts III and IV. For example, if Paolo Palmieri is correct in his re-telling of the history of Galileo's falling bodies thought experiment, this might impact those philosophical accounts that rely heavily on the features of this thought experiment as evidence (perhaps including McAllister 1996, this volume; Norton 1996; Gendler 1998 and Brown 2011). And the accounts presented in Parts III and IV will inspire new ways to interpret the examples considered in Parts I and II. To this end, we have encouraged as much interaction between the chapters of the different parts of this book as possible, and we hope this has improved the collection.

Of course, fruitful interaction between history, philosophy and social and cognitive science has been present in the literature on thought experiments since its beginning. Accordingly, most chapters in this companion present historical arguments with epistemological and metaphysical assumptions in mind, or philosophical arguments that depend on specific readings of history, cognitive science, or social psychology. Thus, while the parts of this book are separated in a way that we hope will make them easy to access and pleasant to read, the distinctions between the

parts are always more or less artificial. For instance, Kenneth Westphal's chapter (Part I) could easily have been in Part IV, and James McAllister's chapter (Part IV) could easily have been in Part I.

One final point: this collection draws together many different sorts of first and second order knowledge, including: (1) expertise in the history and methods of the disciplines discussed in the chapters (theology, history, philosophy, physics, biology, etc.), and (2) expertise in the fields from which the analyses are carried out, that is: history, sociology, cognitive science and philosophy as they are currently practiced. No one of the three editors of this companion pretends to possess such comprehensive knowledge. But rather than cater the chapters to what an average reader could digest, we've chosen to include a wide range of chapters all of which rely to some degree on different background knowledge, in order to maximize breadth. For example, Lenhard assumes familiarity with Monte Carlo simulations and Markov chains, while Palmieri assumes some Husserl. We hope the reader finds this exciting rather than discouraging.

The topic of thought experiments – perhaps because it draws together so many disciplines – appears boundless. Though much has been covered, there is very much more to be done. As at least one of us is inclined to say: Welcome to Plato's heaven.

Notes

1   Please see also our special index cataloguing the thought experiments discussed in this Companion (page 558).

2   For Searle's Chinese room, see Arthur (this volume). For Rawls' original position, see Miščević (this volume), and Brun (this volume). For Plato's thought experiments, see Becker (this volume). For the trolley problem, see Brun (this volume). For more philosophical thought experiments, see Dennett (2013).

3   For Lucretius' thought experiment, see Brown (2011), Ierodiakonou (2011), Brown and Stuart (2013); Meynell (this volume). For Maxwell's demon, see Brown (2011), Buzzoni (2008, 97–100), Krimsky (1973), Myrvold (2011), Norton (this volume), Peacocke (this volume), Schlesinger (1996, 473–76), Stuart (2016a). For Einstein's thought experiments, see Norton (1991) Peacocke (this volume), Meynell (this volume). For Schrödinger's cat, see Peacocke (this volume). For Newton's thought experiments, see Arthur (this volume), Norton (1996); Peijnenburg and Atkinson (2003); Brown (2011). For Heisenberg's microscope, see Popper (1959), Stuart (2016a), Van Dyck (2003), and Camilleri (2007). For Galileo's thought experiments, see Brown (2011), Norton (1996), Gendler (1998, 2000), Buzzoni (2008, 106–7),

Palmieri (2003, this volume), McAllister (1996, this volume). For Stevin's chain, see Brown (2011), Norton (1996, 349–51), and Mach (1905). For thought experiments in mathematics, see Buzzoni (2011), Witt-Hansen (1976), Müller (1969), Brown (1999, 2007a/b), Glas (1999), Van Bendegem (2003), Sherry (2006), Starikova (2007), Starikova and Giaquinto (this volume) and Cohnitz (2008). For thought experiments in geometry, see Lakatos (1976). In infinity, see Galilei (1638, 32), and Hilbert (2013).

4   See Antsey and Vanzo (2016) for a look at the relationship between experimental philosophy and thought experiments going back to the seventeenth century. See also Systema and Buckwalter (2016), Stich and Tobia (this volume), Ludwig (this volume) and Cohnitz and Häggqvist (this volume).

5   This wording is pretty close to the way Kuhn framed the problem, although not exactly. In Kuhn's words the problem is: "How, then, relying exclusively upon familiar data, can a thought experiment lead to new knowledge or to a new understanding of nature?" (1977, 241).

6   E.g., Aligica and Evans (2009), Arthur (1999), Bishop (1998, 1999), Bokulich (2001), Brendel (2004), Brown (1991a, 1992, 2004, 2007a), Butkovic (2007), Buzzoni (2008), Camilleri (2014), Chandrasekharan, Nersessian and Subramanian (2013), Clatterbuck (2013), Cooper (2005), Davies (2007), De Baere (2003), De Mey (2003, 2005, 2006a, 2006b), Ducheyne (2006), Fehige (2012, 2013), Gendler (1998, 2000, 2004), Georgiou (2007), Gooding (1992, 1994), Häggqvist (1996, 2007, 2009), Hopp (2014), Horowitz and Massey (1991), Humphreys (1993), Irvine (1991), Kujundzic (1998), Laymon (1991), Machery (2011), McAllister (1996), McComb (2013), Moue, Masavetas, and Karayianni (2006), Nersessian (1992, 2007). Norton (1991, 1996, 2004a), Pitcha (2011), Roberts (1993), Schlesinger (1996), Shepard (2008). Sorensen (1992), Urbaniak (2012) and Wilson (1991).

7   For other criticisms of Norton's view, see e.g., Gendler (1998) and Stuart (2016b).

8   See Brendel (this volume) for criticism of Häggqvist's use of this argument against Norton.

9   More needs to be said here: Brown does portray the intuitive aspect of a thought experiment as something akin to sense experience, though he does not agree that he is committed to there being only one kind of epistemological account for both sense experience and rational insight.

10   Examples of pluralists might be Bokulich and Frappier (this volume), Cohnitz and Häggqvist (this volume) and Stich and Tobia (this volume), while Rachel Cooper (2005) argues against pluralism. An example of a contextualist might be McAllister (this volume). Norton (2004a and 2004b) argues against this approach, since for him the mark of a good thought experiment "cannot be something external to the thought experiment; that is, something about

the person who authors the thought experiment or about the context in which it is proposed" (2004b, 54). Almost all who write on thought experiments are skeptical concerning at least some of their uses, but it is difficult to find anyone who is skeptical about thought experiments tout court. For example, Pierre Duhem is often quoted as a skeptic, but Buzzoni argues convincingly that he was only drawing attention to some of the dangers we face when putting too much weight on them (Buzzoni forthcoming). Daniel Dennett has been cited as a skeptic by many, and perhaps for good reason: he calls thought experiments "intuitions pumps," and admits that his first use of this term "was derogatory" (1996, 182). But again, his skepticism was only directed toward the over-confident use of thought experiments. In later work, he writes,

> If you look at the history of philosophy, you see that all the great and influential stuff has been technically full of holes but utterly memorable and vivid. They are what I call 'intuition pumps' – lovely thought experiments. Like Plato's cave, and Descartes's evil demon, and Hobbes' vision of the state of nature and the social contract, and even Kant's idea of the categorical imperative. I don't know of any philosopher who thinks any one of those is a logically sound argument for anything. But they're wonderful imagination grabbers, jungle gyms for the imagination. They structure the way you think about a problem. These are the real legacy of the history of philosophy. A lot of philosophers have forgotten that, but I like to make intuition pumps. (Dennett 1996, 182)

Paul Thagard is another candidate for a skeptic. He argues that "the made-up thought experiments favored by many philosophers are not evidence at all" (2010, 209). Rather, "philosophical attempts to establish truths by a priori reasoning, thought experiments, or conceptual analysis have been no more successful than faith-based thinking has been. All these methods serve merely to reinforce existing prejudices" (2010, 41). Still, Thagard allows thought experiments a function in science, and his naturalism about philosophy implies that the same methods should be used there as in science, which a fortiori grants a role for thought experiments in philosophy (see Stuart 2014. See also Buzzoni 2016; and Fehige, this volume). The best candidate for a true skeptic might be Alexius Meinong, who argues against Mach that "an experiment that in fact does not exist at all, can neither prove nor teach anything" (1907, 276–77).

11   This question can also be thought of as an instantiation of Stich and Tobia's main question, given that modal claims are one of the extra-mental entities they discuss.

References

Aligica, P. D. and Evans, A. J. (2009) "Thought experiments, counterfactuals and comparative analysis," The Review of Austrian Economics 22: 225–239.

Antsey, P. and Vanzo, A. (2016) "Early modern experimental philosophy," in A Companion to Experimental Philosophy, edited by J. Systema and W. Buckwalter, Somerset, UK: Wiley-Blackwell.

Arthur, R. (1999) "On thought experiments as a priori science," International Studies in the Philosophy of Science 13: 215–229.

Bishop, M. (1998) "An epistemological role for thought experiments," in Idealization IX: Idealization in Contemporary Physics, edited by N. Shanks, Amsterdam: Rodopoi.

Bishop, M. (1999) "Why thought experiments are not arguments," Philosophy of Science 66: 534–541.

Bokulich, A. (2001) "Rethinking thought experiments," Perspectives on Science 9: 285–307. Brendel, E. (2004) "Intuition pumps and the proper use of thought experiments," Dialectica 58: 89–108.

Brown, J. R. (1986) "Thought experiments since the scientific revolution," International Studies in the Philosophy of Science 1: 1–15.

Brown, J. R. (1991a [2011]) The Laboratory of the Mind: Thought Experiments in the Natural Sciences, London: Routledge.

Brown, J. R. (1991b) "Thought experiments: A Platonic account," in Thought Experiments in Science and Philosophy, edited by T. Horowitz and G. Massey, Lanham: Rowman and Littlefield.

Brown, J. R. (1992) "Why empiricism won't work," Proceedings of the Philosophy of Science Association 2: 271–279.

Brown J. R. (1999) Philosophy of Mathematics: An Introduction to the World of Proofs and Pictures, London: Routledge.

Brown, J. R. (2004) "Why thought experiments do transcend empiricism," in Contemporary Debates in the Philosophy of Science, edited by C. Hitchcock, Malden: Blackwell.

Brown, J. R. (2007a) "Thought experiments in science, philosophy, and mathematics," Croatian Journal of Philosophy 7: 3–27.

Brown, J. R. (2007b) "Comments and replies," Croatian Journal of Philosophy 7: 249–268.

Brown, J. R. and Stuart, M. T. (2013) "Review: Ierodiakonou, K. and Roux, S. (eds), Thought Experiments in Methodological and Historical Contexts," HOPOS: The Journal of the International Society for the History of Philosophy of Science 3: 154–157.

Butkovic, A. (2007) "What is the function of thought experiments: Kuhn vs. Brown," Croatian Journal of Philosophy VII: 63–67.

Buzzoni, M. (2008) Thought Experiment in the Natural Sciences, Würzburg: Königshausen and Neumann.

Buzzoni, M. (2010) "Empirical thought experiments: A transcendental-operational view," Epistemologia XXXIII: 5–26.

Buzzoni, M. (2011) "On mathematical thought experiments," Epistemologia XXXIV: 61–88. Buzzoni, M. (2013a) "Thought experiments from a Kantian point of view," in Thought Experiments in Science, Philosophy, and the Arts, edited by M. Frappier, L. Meynell, and J. R. Brown, London: Routledge.

Buzzoni, M. (2013b) "On thought experiments and the Kantian a priori in the natural sciences: A reply to Yiftach J. H. Fehige," Epistemologia 2: 277–293.

Buzzoni, M. (2015) "Causality, teleology, and thought experiments in biology," Journal for General Philosophy of Science 46: 279–299. DOI: 10.1007/s10838-015-9293-9.

Buzzoni, M. (2016) "Thought experiments in philosophy: A neo-Kantian and experimentalist point of view," Topoi. DOI: 10.1007/s11245-016-9436-6.

Buzzoni, M. (forthcoming) "Duhem and Mach on thought experiments. Or: Did Duhem really reject Mach's thought experiments?", in Vienna Circle Institute Yearbook, Dordrecht: Springer.

Camilleri, K. (2007) "Indeterminacy and the limits of classical concepts: The transformation of Heisenberg's thought," Perspectives on Science 15: 178–201.

Camilleri, K. (2014) "Toward a constructivist epistemology of thought experiments in science," Synthese 191: 1697–1716.

Camp, E. (2009) "Two varieties of literary imagination: Metaphor, fiction, and thought experiments," Midwest Studies in Philosophy XXXIII: 107–130.

Carroll, N. (2002) "The wheel of virtue: Art, literature, and moral knowledge," Journal of Aesthetics and Art Criticism 60: 3–26.

Chandrasekharan, S., Nersessian, N. and Subramanian, V. (2013) "Computational modeling: Is this the end of thought experiments in science?", in Thought Experiments in Science, Philosophy, and the Arts, edited by M. Frappier, L. Meynell, and J. R. Brown, London: Routledge.

Clatterbuck, H. (2013) "The epistemology of thought experiments: A non-eliminativist, non-Platonic account," European Journal for Philosophy of Science 3: 309–329.

Cohen, M. (2004) Wittgenstein's Beetle and Other Thought Experiments, Somerset: Wiley-Blackwell.

Cohnitz, D. (2008) "Ørsteds 'Gedankenexperiment': Eine Kantianische Fundierung der Infinitesimalrechnung? Ein Beitrag zur Begriffsgeschichte von 'Gedankenexperiment' und zur Mathematikgeschichte des frühen 19. Jahrhunderts," Kant-Studien 99: 407–433.

Cooper, R. (2005) "Thought experiments," Metaphilosophy 36: 328–347.

Daiber, J. (2001) Experimentalphysik des Geistes: Novalis und das romantische Experiment, Göttingen: Vadenhoeck and Ruprecht.

Davies, D. (2007) "Thought experiments and fictional narratives," Croatian Journal of Philosophy VII: 29–45.

Davies, D. (2012) "Can philosophical thought experiments be 'screened'?", in Thought Experiments in Science, Philosophy, and the Arts, edited by M. Frappier, L. Meynell, and J. R. Brown, London: Routledge.

De Baere, B. (2003) "Thought experiments, rhetoric, and possible worlds," Philosophica 72: 105–130.

De Mey, T. (2003) "The dual nature view of thought experiments," Philosophica 72: 61–78. De Mey, T. (2005) "Remodeling the past," Foundations of Science 10: 47–66.

De Mey, T. (2006a) "Kuhn's paradox of thought experiments resolved," The Baltic International Yearbook of Cognition, Logic and Communication 1: 111–125.

De Mey, T. (2006b) "Imagination's grip on science," Metaphilosophy 37: 222–239.

Dennett, D. (1996) "Intuition pumps," in Third Culture: Beyond the Scientific Revolution, edited by J. Brockman, New York: Simon and Schuster.

Dennett, D. (2013) Intuition Pumps and Other Tools for Thinking, New York: W. W. Norton and Company.

Ducheyne, S. (2006) "The argument(s) for universal gravitation," Foundations of Science 11: 419–447.

Elgin, C. Z. (1993) "Understanding art and science," Synthese 95: 13–28.

Elgin, C. Z. (2000) "Interpretation and understanding," Erkenntnis 52: 175–183.

Elgin, C. Z. (2002) "Creation as reconfiguration," International Studies in the Philosophy of Science 16: 13–25.

Elgin, C. Z. (2004) "True enough," Philosophical Issues 14: 113–121.

Elgin, C. Z. (2006) "From knowledge to understanding," in Epistemology Futures, edited by S. Hetherington, Oxford: Clarendon Press.

Elgin, C. Z. (2014) "Fiction as thought experiment," Perspectives on Science 22: 221–241.

Fehige, Y. (2012) "'Experiments of pure reason': Kantianism and thought experiments in science," Epistemologia XXXV: 141–160.

Fehige, Y. (2013) "The relativized a priori and the laboratory of the mind: Towards a neo-Kantian account of thought experiments in science," Epistemologia XXXVI: 55–73.

Fehige, Y. and Stuart, M. T. (2014) "On the origins of the philosophy of thought experiments: The forerun," Perspectives on Science 22: 179–220.

Fehige, Y. and Wiltsche, H. A. (2013): "The body, thought experiments, and phenomenology," in Thought Experiments in Science, Philosophy, and the Arts, edited by M. Frappier, L. Meynell, and J. R. Brown, London: Routledge.

Froese, T. and Gallagher, S. (2010) "Phenomenology and artificial life: Toward a technological supplementation of phenomenological methodology," Husserl Studies 26: 83–106.

Galilei, G. (1638 [1954]) Dialogues Concerning Two New Sciences, translated by H. Crew and A. de Salvio, New York: Dover.

Gellard, C. (2011) "Thought experiments in late Medieval debates on atomism," in Thought Experiments in Methodological and Historical Contexts, edited by K. Ierodiakonou and S. Roux, Leiden: Brill.

Gendler, T. S. (1998) "Galileo and the indispensability of scientific thought experiment," British Journal for the Philosophy of Science 49: 397–424.

Gendler, T. S. (2000) Thought Experiment: On the Powers and Limits of Imaginary Cases, London: Routledge.

Gendler, T. S. (2004) "Thought experiments rethought—and reperceived," Philosophy of Science 71: 1152–1163.

Gendler, T. S. and Hawthorne, J. (2002) Conceivability and Possibility, Oxford: Oxford University Press.

Georgiou, A. (2007) "An embodied cognition view of imagery-based reasoning in science: Lessons from thought experiments," Croatian Journal of Philosophy VII: 29–45.

Glas, E. (1999) "Thought-experimentation and mathematical innovation," Studies in History and Philosophy of Science 30: 1–19.

Gooding, D. (1992) "What is experimental about thought experiments?," PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 2: 280–290.

Gooding, D. (1994) "Imaginary science," The British Journal for the Philosophy of Science 45: 1029–1045.

Hacking, I. (1992) "Do thought experiments have a life of their own? Comments on James Brown, Nancy Nersessian and David Gooding," PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 2: 302–308.

Häggqvist, S. (1996) Thought Experiments in Philosophy, Stockholm: Almqvist and Wiksell International.

Häggqvist, S. (2007) "The a priori thesis: A critical assessment," Croatian Journal of Philosophy VII: 47–61.

Häggqvist, S. (2009) "A model for thought experiments," Canadian Journal of Philosophy 39: 55–76.

Hilbert, D. (2013) David Hilbert's Lectures on the Foundations of Arithmetics and Logic 1917–1933, edited by W. Ewald and W. Sieg, Heidelberg: Springer-Verlag.

Hopp, W. (2014) "Experiments in thought," Perspectives on Science 22: 76–97.

Horowitz, T. and Massey, G. (eds) (1991) Thought Experiments in Science and Philosophy, Lanham: Rowman and Littlefield.

Humphreys, P. (1993) "Seven theses on thought experiments," in Philosophical Problems on the Internal and External Worlds: Essays on the Philosophy of Adolf Grunbaum, edited by J. Earman, A. I. Janis, G. J. Massey, and N. Rescher, Pittsburgh: University of Pittsburgh Press.

Ichikawa, J. and Jarvis, B. (2009) "Thought-experiment intuitions and truth in fiction," Philosophical Studies 142: 221–246.

Ierodiakonou, K. (2005) "Ancient thought experiments: A first approach," Ancient Philosophy 25: 125–140.

Ierodiakonou, K. (2011) "Remarks on the history of an ancient thought experiment," in Thought Experiments in Methodological and Historical Contexts, edited by K. Ierodiakonou and S. Roux, Leiden: Brill.

Irvine, A. D. (1991) "Thought experiments in scientific reasoning," in Thought Experiments in Science and Philosophy, edited by T. Horowitz and G. Massey, Lanham: Rowman and Littlefield.

Knuuttila, S. and Kukkonen, T. (2011) "Thought experiments and indirect proofs in Averroes, Aquinas, and Buridan," in Thought Experiments in Methodological and Historical Contexts, edited by K. Ierodiakonou and S. Roux, Leiden: Brill.

Krimsky, S. (1973) "The use and misuse of critical Gedankenexperimente," Zeitschrift für allgemeine Wissenschaftstheorie 4: 323–334.

Kuhn, T. S. (1977) "A function for thought experiments," in The Essential Tension, Chicago: University of Chicago Press.

Kühne, U. (2005) Die Methode des Gedankenexperiments, Frankfurt: Suhrkamp.

Kujundzic, N. (1995) "Thought experiments: Architecture and economy of thought," Journal of the British Society of Phenomenology 26: 86–93.

Kujundzic, N. (1998) "The role of variation in thought experiments," International Studies in the Philosophy of Science 12: 239–243.

Lakatos, I. (1976) Proofs and Refutations, Cambridge: Cambridge University Press.

Lautner, P. (2011) "Thought experiments in the De Anima commentaries," in Thought Experiments in Methodological and Historical Contexts, edited by K. Ierodiakonou and S. Roux, Leiden: Brill.

Laymon, R. (1991) "Thought experiments by Stevin, Mach and Gouy: Thought experiments as ideal limits and as semantic domains," in Thought Experiments in Science and Philosophy, edited by T. Horowitz and G. Massey, Lanham: Rowman and Littlefield.

Lennox, J. G. (1991) "Darwinian thought experiments: A function for just-so stories," in Thought Experiments in Science and Philosophy, edited by T. Horowitz and G. Massey, Lanham: Rowman and Littlefield.

Mach, E. (1905) "On thought experiments," in Knowledge and Error: Sketches on the Psychology of Enquiry, translated by T. J. McCormack, Dordrecht: D. Reidel Publishing Company.

Machery, E. (2011) "Thought experiments and philosophical knowledge," Metaphilosophy 42: 191–214.

McAllister, J. (1996) "The evidential significance of thought experiment in science," Studies in the History of the Philosophy of Science 27: 233–250.

McComb, G. (2013) "Thought experiment, definition, and literary fiction," in Thought Experiments in Science, Philosophy, and the Arts, edited by M. Frappier, L. Meynell, and J. R. Brown, London: Routledge.

Meinong, A. (1907 [1973]) "Das Gedankenexperiment," in Über die Stellung der Gegenstandstheorie im System der Wissenschaften, edited by R. Haller and R. Kindinger, Graz-Austria: Akademische Druck und Verlagsanstalt.

Meynell, L. (2014) "Imagination and insight: A new account of the content of thought experiments," Synthese 191: 4149–4168.

Miščević, N. (1992) "Mental models and thought experiments," International Studies in the Philosophy of Science 6: 215–226.

Miščević, N. (2004) "The explainability of intuitions," Dialectica 58: 43–70.

Miščević, N. (2007) "Modelling intuitions and thought experiments," Croatian Journal of Philosophy VII: 181–214.

Mohanty, J. N. (1991) "The method of imaginative variation in phenomenology," in Thought Experiments in Science and Philosophy, edited by T. Horowitz and G. Massey, Lanham: Rowman and Littlefield.

Moue, A., Masavetas, K. and Karayianni, H. (2006) "Tracing the development of thought experiments in the philosophy of the natural sciences," Journal for General Philosophy of Science 37: 61–75.

Müller, I. (1969) "Euclid's Elements and the axiomatic method," The British Journal for the Philosophy of Science 20: 289–309.

Myrvold, W. (2011) "Statistical mechanics and thermodynamics: A Maxwellian view," Studies in History and Philosophy of Science Part B 42: 237–243.

Nersessian, N. J. (1992a) "How do scientists think? Capturing the dynamics of conceptual change in science," in Cognitive Models of Science, edited by R. N. Giere, Minneapolis: University of Minnesota Press.

Nersessian, N. J. (1992b) "In the theoretician's laboratory: Thought experimenting as mental modeling," Proceedings of the Philosophy of Science Association 2: 291–301.

Nersessian, N. J. (2007) "Thought experiments as mental modelling: Empiricism without logic," Croatian Journal of Philosophy VII: 125–161.

Nersessian, N. J. (2008) Creating Scientific Concepts, Cambridge: MIT Press.

Norton, J. D. (1991) "Thought experiments in Einstein's work," in Thought Experiments in Science and Philosophy, edited by T. Horowitz and G. Massey, Lanham: Rowman and Littlefield.

Norton, J. D. (1996) "Are thought experiments just what you thought?," Canadian Journal of Philosophy 26: 333–366.

Norton, J. D. (2004a) "On thought experiments: Is there more to the argument?," Philosophy of Science 71: 1139–1151.

Norton, J. D. (2004b) "Why thought experiments do not transcend empiricism," in Contemporary Debates in the Philosophy of Science, edited by C. Hitchcock, Malden: Blackwell.

Ørsted, H. C. (1811 [1920]) "Første Indledning til den almindelige Naturlaere," in Hans Christian Ørsted: Naturvidenskabelige Skrifter: Samlet Udgave me to Afhandlinger om Hans Virke, vol. 3, edited by K. Meier, Kopenhagen: Andr. Fred. Høst and Søn.

Palmerino, C. R. (2011) "Galileo's use of Medieval thought experiments," in Thought Experiments in Methodological and Historical Contexts, edited by K. Ierodiakonou and S. Roux, Leiden: Brill.

Palmieri, P. (2003) "Mental models in Galileo's early mathematization of nature," Studies in History and Philosophy of Science 34: 229–264.

Peijnenburg, J. and Atkinson, D. (2003) "When are thought experiments poor ones?," Journal for General Philosophy of Science 34: 305–322.

Pitcha, M. (2011) "How to reconstruct a thought experiment," Organon F 18: 154–188.

Popper, K. (1959) "On the use and misuse of imaginary experiments, especially in quantum theory," in The Logic of Scientific Discovery, London: Hutchinson.

Radder, H. (1996) In and About the World: Philosophical Studies of Science and Technology, Albany: State University of New York Press.

Rescher, N. (1977) Dialectics: A Controversy-Oriented Approach to the Theory of Knowledge, Albany: SUNY Press.

Roberts, F. (1993) "Thought experiments and social transformation," Journal for the Theory of Social Behaviour 23: 399–421.

Salis, F. and Frigg, R. (forthcoming) "Capturing the scientific imagination," in The Scientific Imagination, edited by P. Godfrey-Smith and A. Levy, New York: Oxford University Press.

Schick, T. and Vaughn, L. (2012) Doing Philosophy: An Introduction through Thought Experiments, 5th ed., New York: McGraw-Hill.

Schildknecht, C. (1990) Philosophische Masken: Literarische Formen der Philosophie bei Platon, Descartes, Wolff und Lichtenberg, Stuttgart: Metzler.

Schlesinger, G. (1996) "The power of thought experiments," Foundations of Physics 26: 467–482. Shepard, R. N. (2008) "The step to rationality: The efficacy of thought experiments in science, ethics, and free will," Cognitive Science 32: 3–35.

Sherry D. (2006) "Mathematical reasoning: Induction, deduction and beyond," Studies in History and Philosophy of Science 37: 489–504.

Sorensen, R. (1992) Thought Experiments, Oxford: Oxford University Press.

Starikova, I. (2007) "Picture-proofs and Platonism," Croatian Journal of Philosophy 7: 81–92. Stuart, M. T. (2014) "Cognitive science and thought experiments: A refutation of Paul Thagard," Perspectives on Science 22: 98–121.

Stuart, M. T. (2015) Thought Experiments in Science, Ph.D. thesis, University of Toronto, Canada.

Stuart, M. T. (2016a) "Taming theory with thought experiments: Understanding and scientific progress," Studies in the History and Philosophy of Science 58: 24–33.

Stuart, M. T. (2016b) "Norton and the logic of thought experiments," Axiomathes 26: 451–466.

Stuart, M. T. (2017) "Imagination: A sine qua non of science," Croatian Journal of Philosophy vol. XVII, No. 49.

Systema, J. and Buckwalter, W. (eds) (2016) A Companion to Experimental Philosophy, Somerset: Wiley-Blackwell.

Tetlock, P. E. and Belkin, A. (eds) (1996) Counterfactual Thought Experiments in World Politics, Princeton: Princeton University Press.

Thagard, P. (2010) The Brain and the Meaning of Life, Princeton: Princeton University Press. Tittle, P. (2004) What If … : Collected Thought Experiments in Philosophy, Essex: Pearson.

Urbaniak, R. (2012) "'Platonic' thought experiments: How on earth?," Synthese 187: 731–752. Van Bendegem, J. P. (2003) "Thought experiments in mathematics: Anything but proof," Philosophica 72: 9–33.

Van Dyck, M. (2003) "The roles of one thought experiment in interpreting quantum mechanics: Werner Heisenberg meets Thomas Kuhn," Philosophica 72: 79–103.

Walton, K. (1990) Mimesis as Make-Believe: On the Foundations of the Representational Arts, Cambridge: Harvard University Press.

Wartenberg, T. E. (2007) Thinking on Screen: Film as Philosophy, London: Routledge.

Williamson, T. (2005) "Armchair philosophy, metaphysical modality and counterfactual thinking," Proceedings of the Aristotelian Society 105: 1–23.

Williamson, T. (2016) "Knowing by imagining," in Knowledge through Imagination, edited by A. Kind and P. Kung, Oxford: Oxford University Press.

Wilson, M. (1991) "Reflections on strings," in Thought Experiments in Science and Philosophy, edited by T. Horowitz and G. Massey, Lanham: Rowman and Littlefield.

Wiltsche, H. A. (2013) "How essential are essential laws? A thought experiment on physical things and their givenness in adumbrations," in Wahrnehmen, Fühlen, Handeln. Phänomenologie im Wettstreit der Methoden, edited by I. Günzler and K. Mertens, Münster: Mentis.

Witt-Hansen, J. (1976) "H.C. Ørsted, Immanuel Kant, and the thought experiment," Danish Yearbook of Philosophy 13: 48–65.

Yourgrau, W. (1962) "On the logical status of so-called thought experiments," in Proceedings of the Tenth International Conference of the History of Science, Paris: Hermann.

Yourgrau, W. (1967) "On models and thought experiments in quantum theory," Monatsberichte der Deutschen Akademie der Wissenschaften zu Berlin 9: 886–874.