Barry Stocker • Michael Mack
Editors

The Palgrave Handbook of Philosophy and Literature

palgrave
macmillan

Editors
Barry Stocker
Istanbul Technical University
Istanbul, Turkey

Michael Mack
Durham University
Durham, UK

ISBN 978-1-137-54793-4 ISBN 978-1-137-54794-1 (eBook)
https://doi.org/10.1057/978-1-137-54794-1

Library of Congress Control Number: 2018953690

Cover illustration: Tomas Abad / Alamy Stock Photo

This Palgrave Macmillan imprint is published by the registered company Springer Nature Limited
The registered company address is: The Campus, 4 Crinan Street, London, N1 9XW, United Kingdom

Contents

Notes on Contributors

Mathew Abbott is a Lecturer in Philosophy at Federation University Australia. Drawing on modern European and post-Wittgensteinian thought, his research is concerned with intersections of aesthetics, politics, and ethics. He is the author of *Abbas Kiarostami and Film-Philosophy* (Edinburgh) and editor of *Michael Fried and Philosophy: Modernism, Intention, and Theatricality* (Routledge).

Bill Ashcroft is a renowned critic and theorist, founding exponent of postcolonial theory, and co-author of *The Empire Writes Back*, the first text to offer a systematic examination of the field of postcolonial studies. He is the author and co-author of 17 books and over 180 articles and chapters, variously translated into six languages, and he is on the editorial boards of ten international journals. His latest work is *Utopianism in Postcolonial Literatures*. He teaches at the University of New South Wales (NSW) and is a Fellow of the Australian Academy of the Humanities.

Tudor Balinisteanu is a Senior Research Fellow in English Literature at University of Suceava. He is the author of *Religion and Aesthetic Experience in Joyce and Yeats* (Palgrave 2015); *Violence, Narrative and Myth in Joyce and Yeats: Subjective Identity and Anarcho-Syndicalist Traditions* (Palgrave 2013); and *Narrative, Social Myth, and Reality in Contemporary Scottish and Irish Women's Writing: Kennedy, Lochhead, Bourke, Ní Dhuibhne, and Carr* (CSP 2009). He also published in a number of UK, Irish, Canadian, and US journals.

Michelle Boulous Walker lectures in Philosophy at The University of Queensland in Australia. She is the author of *Slow Philosophy: Reading Against the Institution* (Bloomsbury 2017) and *Philosophy and the Maternal Body: Reading Silence* (Routledge 1998). She is interested in the intersections of philosophy, literature, and film, and her research is on laughter.

Marsha S. Collins is a Professor of Comparative Literature and Royster Professor for Graduate Education at the University of North Carolina at Chapel Hill. She

27

Thought Experiments at the Edge of Conceptual Breakdown

İlhan İnan

Though not every piece of fiction is a thought experiment, a thought experiment can be taken to be a piece of fiction. There is always a story in a thought experiment, one that is merely imaginary, usually contrary to fact, or in the philosopher's jargon a "counterfactual" story. Such a story however significantly differs from a literary one in terms of the purpose of its creation. Unlike a literary author, the primary goal of the creator of a thought experiment is not to entertain the reader, nor is it to produce a form of written art—though achieving such qualities could be a plus. Just like a real scientific experiment conducted in the physical world, a thought experiment is typically designed to test a hypothesis and is therefore intended to function as an epistemic tool to expand our knowledge on a particular topic. It differs however from a physical experiment in that it is conducted not in a laboratory or an observatory, but within the human mind—hence its name. The hypothesis that is under test within a thought experiment is usually one that cannot be settled by empirical means. If there was an easy way to construct a physical experiment to test such a hypothesis, there would have been no need for a thought experiment. One reason why a physical experiment may not be a viable option is because the hypothesis to be tested is one that is an a priori claim, rather than an empirical one. Given that philosophical hypotheses are generally regarded as being a priori claims, thought experiments have been widely utilized as instruments to put them to test. This does not imply that a thought

İ. İnan (✉)
Koç University, Istanbul, Turkey
e-mail: iinan@ku.edu.tr

B. Stocker, M. Mack (eds.), *The Palgrave Handbook of Philosophy and Literature*,
https://doi.org/10.1057/978-1-137-54794-1_27

experiment always has a priori content and is thus not about physical reality. Scientists, especially physicists, have made extensive use of thought experiments to test empirical hypotheses and continue to do so. For instance, in order to argue against the Aristotelian idea that heavy objects fall faster than lighter ones, Galileo came up with an ingenious thought experiment, in which he asked us to imagine two balls having different weights tied to each other by a rope. Then we consider the question: if we let go of the tied balls from the top of the Pisa tower, will it fall faster or slower than the heavy ball itself? If we assume that heavier objects fall faster than lighter ones, then we will have to conclude that the lighter ball will slow down the heavier ball and the whole system will fall slower than the heavy ball all by itself. But then given that the whole system is heavier than the heavy ball itself, we would then have to conclude that it should fall faster. Given this contradiction, we thereby show that Aristotle was mistaken. Though the hypothesis that is being tested is not an a priori claim, Galileo's argument based on his thought experiment can be taken to be an a priori argument in that it is aimed at revealing the internal inconsistency of the Aristotelian view. This is one way in which a thought experiment could be a useful tool not just for philosophy, but for science as well. In fact the concept of a thought experiment was first introduced not by a philosopher, but a physicist. Ernst Mach is usually credited for first coining the German term "Gedankenexperimente", though some trace the origins of the development of the concept to the works of earlier physicists.[1] Though the term is relatively new, the employment of thought experiments goes back to at least Ancient Greek philosophy. Plato makes extensive use of them in his famous dialogues. In the *Republic*, for instance, Plato raises the question: is it always good to be moral? While Socrates wishes to argue that the answer ought to be positive, Glaucon raises a challenge that makes use of a famous thought experiment, generally referred to as "the Ring of Gyges". In the fictitious story, we are asked to imagine that Gyges comes across a ring that makes him invisible whenever he wears it. This gives him the power to commit immoral acts such as robbery and rape, without being caught. Is there any reason for Gyges not to commit such acts? This is the question that goes with the story, and it is what makes the story a thought experiment. Glaucon's challenge forces Socrates to construct an argument to show that we have reason to be moral even when we have no fear of punishment. Not all famous Ancient Greek thought experiments had to do with morality; some had to do with physics. For instance, Zeno of Elea's "paradoxes of motion" all involved stories, the most famous of which concerns a race between the athlete Achilles and a tortoise who is given a short head start. We are asked if Achilles will finally catch up with the tortoise and win the race. Zeno gives a brilliant

argument that he can't, from which he concluded that motion is an illusion, which is what gives the paradoxical nature to the story.

Contrary to a literary story, a thought experiment is always aimed at answering a question, which is what gives its experimental character. We may even wish to identify a thought experiment with a story/question pair. This however would be putting it rather crudely; first because in order for a story/question pair to serve as a thought experiment, its question must be generalizable. Given that a story always involves an element of specificity, the question that goes with the story will also be specific, that is, it will be about specific people, places, or objects. Galileo's thought experiment that we just considered is specifically about two imaginary balls, from which we generalize to a universal question concerning motion: is acceleration dependent on mass? Second, a thought experiment at times will be more complex containing not one but two or more stories, aimed at raising not a single but a set of questions. For instance, a famous thought experiment that has generated a lot of discussion in contemporary ethics is the so-called Trolley Problem. The original story about the trolley which is due to Philippa Foot was then used by Judith Jarvis Thomson as a part of a more complex thought experiment that involves not one but two separate stories. Here is Thomson's version of the first story:

> Suppose you are the driver of a trolley. The trolley rounds a bend, and there come into view ahead five track workmen, who have been repairing the track. The track goes through a bit of a valley at that point, and the sides are steep, so you must stop the trolley if you are to avoid running the five men down. You step on the brakes, but alas they don't work. Now you suddenly see a spur of track leading off to the right. You can turn the trolley onto it, and thus save the five men on the straight track ahead. Unfortunately, Mrs. Foot has arranged that there is one track workman on that spur of track. He can no more get off the track in time than the five can, so you will kill him if you turn the trolley onto him. Is it morally permissible for you to turn the trolley? (Thomson 1985, 1395–96)

As Thomson notes, we are all inclined to answer this question in the positive. But then we have the second story:

> Now consider a second hypothetical case. This time you are to imagine yourself to be a surgeon, a truly great surgeon. Among other things you do, you transplant organs, and you are such a great surgeon that the organs you transplant always take. At the moment you have five patients who need organs. Two need one lung each, two need a kidney each, and the fifth needs a heart. If they do

> not get those organs today, they will all die; if you find organs for them today, you can transplant the organs and they will all live. But where to find the lungs, the kidneys, and the heart? The time is almost up when a report is brought to you that a young man who has just come into your clinic for his yearly check-up has exactly the right blood-type, and is in excellent health. Lo, you have a possible donor. All you need do is cut him up and distribute his parts among the five who need them. You ask, but he says, "Sorry. I deeply sympathize, but no." Would it be morally permissible for you to operate anyway? (Thomson 1985, 1395–96)

This time almost everyone answers negatively. The question of whether we are contradicting ourselves in our answers to the two cases turns out to be a very important one that has to do with the philosophical dispute concerning whether there is a moral difference between killing someone as opposed to merely letting them die, as well as the more general problem concerning whether it is ever morally permissible to kill an innocent person. What gives this thought experiment its philosophical value is that its question can further be generalized cutting deep into foundational issues in ethics and morality on the dispute between consequentialist theories, such as utilitarianism, and deontological theories such as Kant's duty ethics.

Perhaps the best way to decide whether a thought experiment is philosophical or scientific, or a hybrid of the two, is to look at the nature of its question. If the specific question within a thought experiment is generalizable to a philosophical question, then we may categorize it as a philosophical thought experiment. To take up another famous example, consider Robert Nozick's Experience Machine:

> Suppose there was an experience machine that would give you any experience you desired. Super-duper neuropsychologists could stimulate your brain so that you would think and feel you were writing a great novel, or making a friend, or reading an interesting book. All the time you would be floating in a tank, with electrodes attached to your brain. Should you plug into this machine for life, preprogramming your life experiences? … Of course, while in the tank you won't know that you're there; you'll think that it's all actually happening … Would you plug in? (Nozick 1999, 42–43)

The specific question that goes with the story is whether we would wish to hook up to such a machine. What gives this question—and consequently the thought experiment—its philosophical character is that in turn it can be generalized to various philosophical questions: is there a difference in value between the internal feeling of satisfying a desire and its actual satisfaction? Is pleasure the only intrinsic value? Does having epistemic contact with reality have intrinsic value? And so on. Given that these are taken to be philosophical questions, the thought experiment is considered to be philosophical. On the other hand, the famous story of Schrödinger's Cat is taken to be a scientific thought experiment given that its specific question "is the cat in the box dead or alive?" generalizes to various important questions in quantum physics:

> One can even set up quite ridiculous cases. A cat is penned up in a steel chamber, along with the following device (which must be secured against direct interference by the cat): in a Geiger counter, there is a tiny bit of radioactive substance, so small, that perhaps in the course of the hour one of the atoms decays, but also, with equal probability, perhaps none; if it happens, the counter tube discharges and through a relay releases a hammer that shatters a small flask of hydrocyanic acid. If one has left this entire system to itself for an hour, one would say that the cat still lives if meanwhile no atom has decayed. The first atomic decay would have poisoned it. The psi-function of the entire system would express this by having in it the living and dead cat (pardon the expression) mixed or smeared out in equal parts. (Schrödinger 1935, tr. Trimmer 1980, 328)

With this thought experiment, Schrödinger wanted to demonstrate the paradoxical character of the Copenhagen interpretation of quantum mechanics by showing that it has the contradictory implication that the cat is both dead and alive.

Let us now put aside the use of thought experiments in science and concentrate on philosophy. Despite the extensive use of thought experiments in philosophy, various figures have given arguments casting doubt on their worth. The most common kind of objection has to do with the reliability of the answer we give to a question in a thought experiment. Given that our answer to such a question is not based on observation, we need to appeal to an internal mental skill, which may be taken to be a disposition, or a capacity, at times referred to as a "mental faculty", or what philosophers in general call an "intuition". The question then arises as to whether an appeal to intuitions is a credible way of doing philosophy. Philosophers have been split up in their response to this question.[2] Saul Kripke, a prominent figure who has made use of thought experiments on many occasions, proclaimed that:

> Of course, some philosophers think that something's having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor of anything, myself. I really don't know, in a way, what more conclusive evidence one can have about anything, ultimately speaking. (Kripke 1980, 42)

In diametrical opposition is Hintikka who holds that the use of thought experiments in which we appeal to intuitions that have no theoretical justification ought to be banned in philosophy:

> Unfortunately, the vast majority of appeals to intuition by contemporary philosophers cannot be conceived as controlled thought experiments nor be justified by recasting them as such. In view of such goings-on, I am tempted to suggest, half-jokingly but only *half*-jokingly that the editors of philosophy journals agree to a moratorium on all papers in which intuitions are appealed to, unless the basis of those appeals is made explicit. (Hintikka 1999, 147)

Interestingly, and perhaps somewhat paradoxically, the idea that thought experiments are not credible argumentative instruments in philosophy given that they appeal to intuitions that are biased can be demonstrated by making use of thought experiments. One way in which this could be done is to take up a thought experiment and present it to groups belonging to different cultures, ethnicities, age groups, genders, educational backgrounds, and so on; if it turns out that the majority of the members of one group answer the question within the thought experiment in one way, and majority of the members of another group answer it in the opposite, we could then demonstrate that intuitions are biased. In fact there have been such surveys conducted by philosophers that have been utilized to argue that thought experiments are not always to be trusted in doing good philosophy. One such attempt concerns the credibility of thought experiments that have been extensively utilized within the theory of reference. This has to do with the ongoing rivalry between the descriptivist theory of reference—different versions of which have been attributed to Frege and Russell—and the causal-historical account—championed by Kripke. In an influential paper, Machery et al. (2004) used the results of their surveys that they conducted on different cultures in order to demonstrate that the thought experiments used by Kripke and others to argue against the descriptivist theory are not credible. They argued that their findings reveal that the responses of Westerners in general indicate that they are inclined toward a causal-historical account, whereas the responses of East Asians appear to favor the descriptivist theory. Many such surveys have been conducted on different thought experiments since then, and this has given rise to an emerging discipline called "experimental philosophy". This method of doing philosophy now has many advocates, as well as many foes. What is important for our purposes here is that if we consider experimental philosophy as "good philosophy", and if we take the arguments utilized by its proponents to have epistemic value, then we ought to conclude that making use of thought

experiments can in fact be a way of doing "good philosophy". After all if philosophers had not created thought experiments, then the surveys used by experimental philosophers could not have taken place. We should then conclude that the findings of experimental philosophy do not show that thought experiments have no use in doing philosophy; rather what they reveal is that thought experiments do not have the power to resolve philosophical disputes, or to demonstrate the truth of at least some philosophical hypotheses, given that our intuitions concerning a certain philosophical issue are highly affected and perhaps determined by our culture, our gender, our socioeconomic status, our education, and so on.[3]

There is however what I believe to be a more interesting way in which a thought experiment may reveal that our intuitions are not to be trusted on a particular topic. These are cases in which the divergence of intuitions takes place not between two separate cultures, but rather within the mind of a single individual. This would require us to utilize a thought experiment that contains not one but a pair of stories posing two separate questions, and if it turns out that the intuitive answers given by an individual to the two questions are inconsistent, we could then conclude that such intuitions are not to be trusted, at least not on whatever the particular issue the stories involve. The point can be demonstrated with a simple non-philosophical example. In the year 1999, a great many people celebrated the new year on the 31st of December, believing that it was the turn of the century. They obviously also believed that a century is a period of 100 years, but what they failed to realize was that the two beliefs were in fact inconsistent. Here is a case in which our "intuition" that a transition from the year 1999 to the year 2000 marks the turn of the century is not to be trusted. The reason is clear: the intuition is simply false. Things usually are not as obvious when it comes to a pair of stories in a philosophical thought experiment which appears to bring out in us conflicting intuitions. In Thomson's case, for instance, it may well be argued that the apparent inconsistency in our answers can be explained away. There are however more pressing cases in which such a route does not seem to suggest itself.

The philosophical thought experiments I now wish to consider raise questions that give us what Wittgenstein called a "mental cramp", or a Socratic "Stingray effect". These involve stories to be followed by what appears to be a simple question that we have hard time dealing with, making it difficult, if not impossible, to settle the issue in one way or another. If the question is in one sense unanswerable, then one may ask what value, if any, such a thought experiment has. If one adopts what appears to be the reasonable view that the primary goal of a thought experiment is to ease our way to find an answer to

an important question, then a thought experiment that raises an unanswerable question should have no value given that it can never achieve its goal. If one nevertheless wishes to insist that a thought experiment whose question is unanswerable still may have value, one should then give up this classical model. It is of course best to discuss this issue by making use of several examples. In what follows I will take up three thought experiments, two of them quite popular, one of them a bit less so. They have two things in common; one is that they are all about the concept of *identity*, and the second is that the central question in each story is mind-boggling, so much so that none of the possible answers appear to be plausible, which forces us to consider whether the question is in fact unanswerable.

The oldest of the three is the famous ancient story concerning the ship of Theseus:

> The ship wherein Theseus and the youth of Athens returned had thirty oars, and was preserved by the Athenians down even to the time of Demetrius Phalereus, for they took away the old planks as they decayed, putting in new and stronger timber in their place, insomuch that this ship became a standing example among the philosophers, for the logical question of things that grow; one side holding that the ship remained the same, and the other contending that it was not the same. (Plutarch (75 A.C.E.), tr. Dryden 1966)

Centuries later Thomas Hobbes extended the thought experiment by adding a second story to it:

> For if, for example, that ship of Theseus, concerning the difference whereof made by continual reparation in taking out the old planks and putting in new, the sophisters of Athens were wont to dispute, were, after all the planks were changed, the same numerical ship it was at the beginning; and if some man had kept the old planks as they were taken out, and by putting them afterwards together in the same order, had again made a ship of them, this, without doubt, had also been the same numerical ship with that which was at the beginning; and so there would have been two ships numerically the same, which is absurd. (Hobbes 1655 Part II, Ch. 11, §7)

In the contemporary metaphysics literature on identity, authors have used this extended form of the thought experiment. First we are asked to imagine that all of the original planks of the ship of Theseus are replaced gradually with new ones, and then we consider the question of whether the resulting ship is the same as the original one. In the second story this time we are asked to imagine that the old planks from the original ship are saved, and after all

planks have been replaced by new ones, we are told that a ship is constructed out of the old planks that looks exactly like the original ship. And then we are asked whether this ship is the same as the original ship of Theseus. If we had been asked this question without first having considered the first story, we could be inclined to answer in the positive. After all there are many artifacts we have observed whose parts are replaced by new ones in time which normally does not imply that the original entity in question ceases to exist and is supplanted by a new entity after a sufficient number of the original entity's parts have been replaced. But once we consider the second story, we hedge. We come to realize that our answers in the two cases jointly will imply that there would be two ships occupying different parts of space, having different spatiotemporal histories that would both be identical to the original ship of Theseus, a conclusion that we have hard time swallowing. We would then have to give up our belief in a principle that appears to be a truism: if x and y are identical to z, then x and y are identical. If we choose not to give it up, the result is even worse, for then we would be forced to hold that two objects located in different parts of space and therefore having different properties can be identical. This is inconsistent with a very intuitive thesis commonly referred to as the Indiscernibility of Identicals which states that if two objects are identical, then they have the same properties. None of the options seem plausible.

A similar thought experiment was constructed by Derek Parfit concerning personal identity:

> I enter the Teletransporter. I have been to Mars before, but only by the old method, a space-ship journey taking several weeks. This machine will send me at the speed of light. I merely have to press the green button. Like others, I am nervous. Will it work? I remind myself what I have been told to expect. When I press the button, I shall lose consciousness, and then wake up at what seems a moment later. In fact I shall have been unconscious for about an hour. The Scanner here on Earth will destroy my brain and body, while recording the exact states of all of my cells. It will then transmit this information by radio. Travelling at the speed of light, the message will take three minutes to reach the Replicator on Mars. This will then create, out of new matter, a brain and body exactly like mine. It will be in this body that I shall wake up. (Parfit 1984, 199)

At this stage we are asked to consider the question whether the person that comes out after the teletransportation is the same person as the original person before the process. Intuitions may diverge, though it seems the majority are willing to answer in the positive. This is, for instance, indicated by the fact

that people who watch Star Trek have no difficulty in believing that it is the same Captain Kirk we see in all the different episodes though he gets beamed to different planets on many occasions. There is however the second story:

> Several years pass, during which I am often Teletransported. I am now back in the cubicle, ready for another trip to Mars. But this time, when I press the green button, I do not lose consciousness. There is a whirring sound, then silence. I leave the cubicle, and say to the attendant: 'It's not working. What did I do wrong?' 'It's working', he replies, handing me a printed card. This reads: 'The New Scanner records your blueprint without destroying your brain and body. We hope that you will welcome the opportunities which this technical advance offers.' The attendant tells me that I am one of the first people to use the New Scanner. He adds that, if I stay for an hour, I can use the Intercom to see and talk to myself on Mars. 'Wait a minute', I reply, 'If I'm here I can't also be on Mars'. Someone politely coughs, a white-coated man who asks to speak to me in private. We go to his office, where he tells me to sit down, and pauses. Then he says: 'I'm afraid that we're having problems with the New Scanner. It records your blueprint just as accurately, as you will see when you talk to yourself on Mars. But it seems to be damaging the cardiac systems which it scans. Judging from the results so far, though you will be quite healthy on Mars, here on Earth you must expect cardiac failure within the next few days.' The attendant later calls me to the Intercom. On the screen I see myself just as I do in the mirror every morning. But there are two differences. On the screen I am not left-right reversed. And, while I stand here speechless, I can see and hear myself, in the studio on Mars, starting to speak. (Parfit 1984, 199–200)

This time we are baffled. The ones who have the intuition that in the first story the person on Mars is the same person as the one before the teletransportation usually are forced to give up their positions after hearing the second story. Otherwise they would have to admit that there are two persons on different planets who are the "same" person. This not only sounds wrong to our ears, but there is a sense in which we find it difficult to comprehend what it actually means. Upon learning that he will soon die of cardiac arrest, the person on Earth should be expected to fear death; but there is some room for conciliation. After all the person on Mars who will survive looks exactly the same and has exactly the same memories and character traits as the person on Earth before the teletransportation takes place. Isn't this enough for him to say "I will survive"? Does such a question have a "correct" answer? If not, then what have we learned from this thought experiment?

The third thought experiment I wish to consider is due to Kripke in his classic piece "A Puzzle about Belief":

> Peter…may learn the name 'Paderewski' "with an identification of the person named as a famous pianist". Naturally, having learnt this, Peter will assent to "Paderewski had musical talent" and we can infer—using 'Paderewski', as we usually do, to name the Polish musician and statesman: …Peter believes that Paderewski had musical talent. Later, in a different circle, Peter learns of someone called 'Paderewski' who was a Polish nationalist leader and Prime Minister. Peter is skeptical of the musical abilities of politicians. He concludes that probably two people, approximate contemporaries no doubt, were both named 'Paderewski'. Using 'Paderewski' as a name for the statesman, Peter assents to 'Paderewski had no musical talent'. (Kripke 1979, 449)

Given the story we now consider the question: does Peter believe that Paderewski had musical talent? There appears to be a strong argument that the answer should be affirmative, but then there also appears to be a strong argument that it should be in the negative. If a speaker sincerely assents to an utterance of a declarative sentence that he or she understands, we normally take that to be sufficient to conclude that the speaker believes the proposition expressed by that sentence. Peter gives his sincere assent to the sentence "Paderewski had musical talent", and it appears that he does grasp what is being said, but then in the second context, Peter this time assents to what appears to be the logical negation of his earlier belief. If we were to ask Peter the question "do you believe that Paderewski had musical talent?", he would say "Yes" in one context, and "No" in the other. From his affirmative answer, we would normally conclude that Peter believes that Paderewski had musical talent; but from his negative answer if we were to conclude that he *fails to believe* that Paderewski had musical talent, we would run into a contradiction.

In all of the three thought experiments, there is a question that appears to arouse in us conflicting intuitions. Is the original ship of Theseus identical to the later one after all of the original planks have been replaced? Am I the same person after being teletransported? Does Peter believe that Paderewski had musical talent? In all three cases, when the question is asked at the end of the first story, it seems that respondents would normally answer in the positive. This is quite obvious in the Paderewski case. If all we are told is that Peter learns that Paderewski was a famous pianist, and assents to the sentence "Paderewski had musical talent", then quite obviously we would normally conclude that Peter believes that Paderewski had musical talent. It is only after we hear the rest of the story, or better, after we hear the second story, that perplexity sets in. Concerning the ship of Theseus, my guess is that a normal adult who speaks a language close to ours would normally say that a ship

remains as the same ship even after all of its parts are replaced gradually. It is only after they hear the second story that they hedge. Now the teletransporter story differs from the previous two in that it involves a bit of science fiction. Unlike the other two cases, the stories involved here may turn out to be impossible in some sense of impossibility. Nonetheless, it seems that people have no problem in conceiving—or at least in believing that they are conceiving—what is being told in the story. For the purposes of our discussion, it really does not matter whether teletransportation is possible or whether it violates some laws of physics. What matters is that the story appears to be one that is easily conceivable, raising questions that appear to be unanswerable. A question may be called an unanswerable question when it actually has a correct answer though it is beyond our capacity to come to know it. "What is it like to be a bat?" We don't know, and we will never come to know, though if there is something it is like to be a bat, then the question does have an answer.[4] For such questions, we may say that there is a fact of the matter that answers the question though it is a fact that is unknowable. The questions we are considering in these thought experiments do not seem to be unanswerable in this sense. It is not that we lack epistemic access to some piece of reality that would answer the question; nor is it that we are not clever and patient enough to reach the correct answer by a long-winded piece of reasoning. It may be that such questions simply do not have correct answers. Such a position immediately raises eyebrows. One reason for this is that philosophers have been reluctant to take an unanswerable question as being meaningful. For instance, Wittgenstein toward the very end of the Tractatus says:

> When the answer cannot be put into words, neither can the question be put into words. The riddle does not exist. If a question can be framed at all, it is also possible to answer it. (Wittgenstein 1961, 6.5)

In the later literature on the logic of questions and answers, this has been the dominant view. An early example of this is due to Knight:

> We have seen that inquiry presupposes…an unknown based on fact….and faith that such knowledge exists….any question which violates one of these presuppositions of inquiry is meaningless for that purpose. (Knight 1967, 571)

It is indeed true that typically when we ask a question, we presuppose that there is an answer which we can at least in principle discover. What happens when we find out that we are mistaken in our presupposition? Well one thing that may happen is that we lose our motivation to find the correct answer. But

why should we all of a sudden have to declare that the question before us is in fact a meaningless question? Before Euclid proved that prime numbers are infinite, it is natural to assume that some people believing that they were finite raised the question "what is the largest prime number?" and tried to find the correct answer. Upon learning that they were mistaken in their presupposition, they must have given up their inquiry. We now know that the question has no answer, though we do not thereby dismiss it as being meaningless. In fact in order to demonstrate that the question has no answer, one first needs to understand what the question is asking.[5]

Going back to our thought experiments, it would be quite preposterous to say that their respective questions are devoid of meaning. Just the opposite, in order for us to be baffled by such questions, first we need to understand what they ask. From this however we cannot conclude that they must have correct answers; it may be the case that there simply is no fact of the matter that answers them. What could that teach us? Well for one thing it would motivate us to try to unearth the implicit presuppositions that have failed us in these contexts. If there is any knowledge to be gained as a result of this, it would have to relate to the limits of our conceptual apparatus. We could thereby learn something, not about the world, but the concepts we use to think about the world. Such conceptual knowledge would of course have to be *negative*, revealing that there are contexts in which our ordinary concepts fail us. If a thought experiment can achieve this, we may say that it has "negative heuristics". This may be the sole value of the three thought experiments we have discussed here. What they show is that our ordinary concept of identity breaks down within these unusual contexts. Such a view appears to have been endorsed by Quine with respect to the concept of personal identity:

> The method of science fiction has its uses in philosophy, but… I wonder whether the limits of the method are properly heeded. To seek what is 'logically required' for sameness of person under unprecedented circumstances is to suggest that words have some logical force beyond what our past needs have invested them with. (Quine 1972, 490)

If a thought experiment could be said to possess "positive heuristics", then it would be a tool that enables us to discover the correct answer to a question, and when it has negative heuristics, then it allows us to discover that the question has no answer.[6] Of course it will always be a matter of controversy which side a particular thought experiment falls under. For instance, in the Paderewski case, contra Kripke, one may wish to argue that the question "Does Peter believe that Paderewski had musical talent?" does have a definite correct

answer. Given that Peter assents to the sentence "Paderewski had musical talent" in one of the contexts in which this is posed to him, some may wish to hold that this should be sufficient for us to conclude that the question should be answered affirmatively. The fact that Peter dissents from the same sentence when posed to him in the other context, they may claim, does not override the earlier belief attribution. The advocates of this view could take Peter's dissent to imply that he disbelieves that Paderewski had musical talent, and not that he fails to believe it. Such a solution to the puzzle results in the attribution of conflicting beliefs to Peter; he simultaneously believes and disbelieves that Paderewski had musical talent. This is certainly not Kripke's position on the matter. At the start of his paper, he explicitly says that his intention is to argue that "the puzzle is a puzzle" and that "(a)ny speculation as to solutions can be deferred". Having said this it may be suggested that the way that Kripke sets up the thought experiment may be taken to support the "solution" mentioned above. When Peter uses "Paderewski" as a name of the statesman, we are told that he assents to the sentence "Paderewski had no musical talent". This of course could be taken to imply that Peter disbelieves that Paderewski had musical talent. We could however slightly modify the example so as not to allow for this. Given that Peter is skeptical of the musical abilities of politicians, we may simply assume that he assents neither to the sentence "Paderewski had musical talent" nor to the sentence "Paderewski had no musical talent" (when he takes the name as the name of the statesman). If so we would not have any reason to conclude that Peter disbelieves that Paderewski had musical talent. Being skeptical does not imply disbelief. We may build it into the story that Peter does not have any particular evidence that Paderewski, when presented to him as the politician, had no musical talent. A background belief that politicians in general do not have musical talent does not provide sufficient epistemic grounds for Peter to believe this. His dissent from the sentence would then indicate, not that he disbelieves the proposition in question, but rather that he fails to believe it, and if so, the puzzle cannot be solved by attributing to Peter conflicting beliefs.[7] Kripke does not go as far as claiming that "the puzzle" has no solution, but comes close to it:

> When we enter into the area exemplified by Jones…, we enter into an area where our normal practices of interpretation and attribution of belief are subjected to the greatest possible strain, perhaps to the point of breakdown. So is the notion of the content of someone's assertion, the proposition it expresses. In the present state of our knowledge, I think it would be foolish to draw any conclusion, positive or negative, about substitutivity. (Kripke 1979, 269)

In fact earlier in his famous *Naming and Necessity* lectures, Kripke had already expressed a similar view:

> My view that the English sentence 'Hesperus is Phosphorus' could sometimes be used to raise an empirical issue while 'Hesperus is Hesperus' could not shows that I do not treat the sentences as completely interchangeable. Further, it indicates that the mode of fixing the reference is relevant to our epistemic attitude toward the sentences expressed. How this relates to the question what 'propositions' are expressed by these sentences, whether these 'propositions' are objects of knowledge and belief, and in general, how to treat names in epistemic contexts, are vexing questions. I have no 'official doctrine' concerning them, and in fact I am unsure that the apparatus of 'propositions' does not break down in this area. Hence, I sidestepped such questions; no firm doctrine regarding the point should be read into my words. (Kripke 1980, 20–21)

Respecting his will I do not wish to attribute to him any official doctrine. In neither of the two passages quoted does Kripke explicitly endorse the view that there is what I have called a "conceptual breakdown" in such contexts, which would make the questions of these thought experiments unanswerable. All that he says is that "perhaps" this may be the case. My concern here is not what Kripke believes, but rather what we can conclude about the use of such thought experiments in philosophy if there is indeed such a conceptual breakdown. In all these cases, there is a question such that when it is posed after the initial story, we are inclined to answer in the positive, but when the very same question is posed under a different scenario, we are inclined to answer in the negative; and once we reflect on our apparent inconsistency, this time we are inclined to enter into the twilight zone where we no longer know what to say on the matter. One reason for this may very well be that a concept that we use daily, such as identity, is not fine-grained enough to be applied to the cases in question.[8] If this is the case, it is exactly what makes these thought experiments valuable and philosophically significant. We may then conclude that some thought experiments have the function of revealing the limits of our conceptual apparatus, which is what gives them negative heuristics. This could then give us the motivation to change our language in a way that would not yield such inconsistencies.[9]

The idea that a thought experiment can have the power to reveal the internal inconsistencies of our conceptual apparatus is true not just of philosophical discourse but is equally applicable to the scientific context as well. Galileo's thought experiment may be taken to reveal the internal inconsistencies of the Aristotelian framework, and the same could be said about Schrödinger's Cat. In fact Thomas Kuhn highlights this function of a thought experiment for science:

> ...the new understanding produced by thought experiments is not an understanding of nature but rather of the scientist's conceptual apparatus. On this analysis, the function of the thought experiment is to assist in the elimination of prior confusion by forcing the scientist to recognize contradictions that had been inherent in his way of thinking from the start. (Kuhn 1977, 242)

If a thought experiment can reveal the "confusions" and "contradictions" of the scientist, it can also do the same for the philosopher. What is more is that when the central concept within a philosophical thought experiment is not merely a part of the technical philosophical jargon, but is a concept that is accessible and understood by anyone who has mastered a language, then coming to realize the confusions and contradictions that are inherent in its use ought to be a concern, not just for the philosopher, but for the layman who uses that concept daily to think and communicate. Such is the case for our concept of identity. Even if the layman does not call it by that name, it is quite obvious that they employ the concept of identity in their daily routines of thinking about their own selves and other people as well as entities such as ships. Not only do they use the concept of identity, but they find themselves in contexts in which it becomes important for them whether it is the same person or the same ship that they are thinking and talking about. One need not be a philosopher to care about such issues. There can, of course, be thought experiments that have to do with not an ordinary concept such as identity, but a technical philosophical notion that only a select few understand and employ. One may, for instance, take Kripke's famous Schmitt case to be a thought experiment concerning the notion of rigid designation, which does not appear to be a part of our daily discourse. This is a technical concept invented by Kripke primarily to argue against a descriptivist theory of reference. No doubt when the thought experiment is introduced by appealing to this notion, the laymen will have difficulty understanding it. This, however, does not imply that the layman does not have intuitions concerning whether proper names are rigid or accidental designators. That is because the story within the thought experiment can be given without having to appeal to this notion. In fact, within the original version of the story, as given by Kripke himself, the notion of rigid designation does not appear at all. In order to understand the story and answer its question, all that one needs is the notion of reference. That is why experimental philosophers have been able to conduct surveys to test the semantic intuitions of speakers from different cultures who have no training in the philosophy of language and have never heard of the notion of rigid designation in their lives. The value of a philosophical thought experiment that reveals our intuitions concerning the way in which we use a specific concept should not be sought in the narrow technical jargon of some philosophical subdiscipline.

The primary and in certain cases the sole value of a thought experiment may be its power in showing us the limits of our conceptual apparatus. Coming up with stories and questions that accomplished this requires creativity. That is why it takes a good philosopher to introduce a good thought experiment. After that it becomes public property, and the non-philosopher could benefit from it. On the classical approach, "the benefit" is identified with expansion of knowledge, and the primary way in which a thought experiment could expand our knowledge is when we come to know the correct answer to its central question, but it appears that in these cases we are unable to achieve this. If these thought experiments have any benefit, it certainly does not seem to fit this classical schema. In the "normal" case, we expect and hope that there will be either a positive or a negative outcome; that is, we will either be in a position to demonstrate the truth of a philosophical hypothesis or to refute it. In the thought experiments we have considered, however, no such outcome is achieved, and this does not appear to be because the problem is too difficult and complicated for us to solve. These are thought experiments that have the Stingray effect. It may very well be the case that they raise questions that are unanswerable, not because of the limits of our epistemic capacity, but because there is no fact of the matter that answers them given that we find ourselves in an area where our conceptual apparatus breaks down. Once it is acknowledged that a thought experiment may have such a role, one may intentionally seek to create thought experiments to achieve this end. Just like in science one may construct what Popper called a "crucial experiment" to test a theory by intending to put the theory to the maximum amount of strain, one may create a thought experiment to do something similar regarding some of our basic concepts. I do not wish to generalize from what I have said concerning these three thought experiments to jump to the sweeping claim that the primary use of thought experiments in philosophy is to reveal deficiencies in our conceptual apparatus within a particular area. That would be too bold. Having said this I should admit that I am not willing to hold that these are marginal cases that have little representative value of thought experiments in philosophy in general.

Acknowledgments I presented some of the ideas in this chapter at Koç University in May 2017. I would like to thank the Philosophy Club, especially its President Selim Utku Öğüt, for the invitation and for being a great host. I am grateful to the audience for the lively discussion and to Kenneth R. Westphal, Erhan Demircioğlu, Zeynep Direk, and Zeynep Talay for their valuable feedback.

Notes

1. See Kühne (2005).
2. There is a wide literature in defense of the utilization of thought experiments in philosophy, for example, Sorensen (1992), Miscevic (2007), Williamson (2008), Thorpe (2016).
3. In a pioneering and influential article in experimental philosophy, Weinberg et al. (2001) try to demonstrate that epistemic intuitions on certain matters can vary significantly across cultures and groups having different socioeconomic status. They make use of certain popular thought experiments that have been utilized by contemporary philosophers (such as the famous Gettier cases), but when there is a need, they make up their own thought experiments as well. If their conclusion is correct, then this would perhaps demonstrate that our raw epistemic intuitions are at times highly affected by our culture. It would not however show that thought experiments are useless vehicles in doing philosophy, given that these authors have made good use of them to arrive at their conclusions.
4. In his classic piece, Thomas Nagel (1974) argues that even if we came to know all the physical properties of bats, we still would not come to know what it is like to be a bat. Though he does not put it in these terms, it is clear that for Nagel the fact that the question is unanswerable for us does not prevent us from grasping what it asks.
5. For a more elaborate discussion on how we may express our curiosity by raising an unanswerable question, see İnan (2012), especially *Chapter 11 Presuppositions of Curiosity* and *Chapter 12 The Limits of Curiosity and Its Satisfaction.*
6. The idea that raising a philosophical question which has no answer can have epistemic merits is not a popular one. To my knowledge, there is no work that directly addresses this issue. Within the wide literature on metaphilosophy, one would at least expect to find some in-depth discussion of the nature of philosophical questions, which unfortunately is not the case. Among the very few works on the topic is a stimulating article titled "What is a Philosophical Question?" published in Mind in 1964 by Nermi Uygur, and a more recent paper by Luciano Floridi (2013) with the same title. Both papers address the issue of "unanswerability" in different ways.
7. Nathan Salmon (1995) offers a solution to the puzzle by claiming that Peter has three separate attitudes to the very same proposition: belief, disbelief, and a third one which he calls "withholding belief". As I have argued in the text given, we may slightly modify the story so that it would be wrong to conclude that Peter has a disbelief in the proposition. Does he withhold belief? If that amounts to failing to believe, we would then have a contradiction. Salmon finds an ingenious way to define "withholding belief" such that it turns out that an agent may both believe a proposition and "withhold belief" from it. It is quite clear to me that such a notion does not correspond to our ordinary use of the notion of *belief.*
8. Machery (2011) argues that within a thought experiment, our "psychological capacity" to make a judgment is "applied outside its proper domain" (p. 201) and thus becomes less credible. It is of course not easy to detect what that psychological capacity is and whether it is the same one in the three cases we have considered. It may be more accurate to say that we cannot even apply our psychological capacity to arrive at a conclusion in these three thought experiments. I should note that Machery too (2004, 2011) (just like Weinberg et al.) makes extensive use of thought experiments in his papers to shed doubt on the credibility of our intuitions, which is another good example of what I have called "the negative heuristics" of thought experiments.
9. There is of course a lot to be said about how a thought experiment can bring about conceptual revision. See Sorensen (1992), chapter 7, for a discussion of this issue.

References

Floridi, Luciano. 2013. What is a Philosophical Question? *Metaphilosophy* 44 (3): 195–221.

Hintikka, Jaakko. 1999. The Emperor's New Intuitions. *Journal of Philosophy* 96 (3): 127–147.

Hobbes, Thomas. 1655. *De Corpere, Elements of Philosophy,* New York: Abaris Books, 1981.

İnan, İlhan. 2012. *The Philosophy of Curiosity.* London: Routledge.

Knight, Thomas S. 1967. Questions and Universals. *Philosophy and Phenomenological Research* 27: 564–576.

Kripke, Saul A. 1979. A Puzzle About Belief. In *Meaning and Use,* ed. A. Margalit, 239–283. Dordrecht: Reidel.

———. 1980. *Naming and Necessity.* Cambridge, MA: Harvard University Press.

Kuhn, Thomas. 1977. *The Essential Tension.* Chicago: The University of Chicago Press.

Kühne, Ulrich. 2005. *Die Methode des Gedankenexperiments.* Frankfurt: Suhrkamp.

Machery, Edouard. 2011. Thought Experiments and Philosophical Knowledge. *Metaphilosophy* 42 (3): 191–214.

Machery, E., R. Mallon, S. Nichols, and S.P. Stich. 2004. Semantics, Cross-Cultural Style. *Cognition* 92 (3): B1–B12.

Miscevic, Nenad. 2007. Modelling Intuitions and Thought Experiments. *Croatian Journal of Philosophy* VII: 181–214.

Nagel, Thomas. 1974. What is It Like to be a Bat? *Philosophical Review* 83 (4): 435–450.

Nozick, Robert. 1999. *Anarchy, State and Utopia* (original published in 1974 by Basic Books) Reprint. Oxford: Blackwell.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
Plutarch (75 A.C.E.). 1966. *Theseus*. Translated by John Dryden. New York, NY: Biblo and Tannen.
Quine, W.V. 1972. Review: Identify and Individuation by Milton K. Munitz. *The Journal of Philosophy* 69 (16): 488–497.
Salmon, Nathan. 1995. Being of Two Minds: Belief with Doubt. *Nous* 29 (1): 1–20.
Schrödinger, Erwin. 1935. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften* 23 (48): 807–812. English translation by John D. Trimmer. 1980. The Present Situation in Quantum Mechanics: A Translation of Schrödinger's 'Cat Paradox Paper'. *Proceedings of the American Philosophical Society* 124 (5): 323–338.
Sorensen, R.A. 1992. *Thought Experiments*. New York, NY: Oxford University Press.
Thomson, Judith Jarvis. 1985. The Trolley Problem. *The Yale Law Journal* 94 (6): 1395–1415.
Thorpe, Lucas. 2016. Experimental Philosophy, Williamson's Expertise Defense of Armchair Philosophy and the Value of the History of Philosophy. *Yeditepe'de Felsefe: Method in Philosophy*, 169–184.
Uygur, Nermi. 1964. What is a Philosophical Question? *Mind* 73 (289): 64–83.
Weinberg, J.M., S. Nichols, and S. Stich. 2001. Normativity and Epistemic Intuitions. *Philosophical Topics* 29 (1–2): 429–460.
Williamson, Timothy. 2008. *The Philosophy of Philosophy*. London: Wiley-Blackwell.
Wittgenstein, Ludwig. 1961. *Tractatus Logico-philosophicus*. Translated by D.F. Pears and B.F. McGuinness. London: Routledge and Kegan Paul.

28

Ethics and Literature

Liesbeth Korthals Altes and Hanna Meretoja

This chapter discusses a number of approaches to the intersections of literature and ethics.[1] Our focus is in literary studies, as we cannot also attempt to cover in sufficiently systematic ways the rich reflection on the topic articulated in literature itself and in philosophy proper. The chapter first provides a brief historical perspective of the kinds of debates waged in relation to literature and ethics. We then zoom in on the following interpretative approaches, characteristic of the so-called ethical turn: criticism inspired by the neo-Aristotelian humanist tradition in moral philosophy; rhetorical criticism, which carries further this humanist tradition, integrating the analysis of narrative form; poststructuralist and deconstructive criticism; social and cultural criticism; and, finally, forms of criticism rooted in philosophical hermeneutics. A marked interest for ethics and literature can also currently be observed in the social sciences. We briefly discuss a number of sociological, cognitive, and psychological approaches that seek to support or qualify claims about literature's ethical potential or position these within broader negotiations of value in culture. Our concluding remarks pertain among others to the question of

L. Korthals Altes (✉)
University of Groningen, Groningen, Netherlands
e-mail: e.j.korthals.altes@rug.nl

H. Meretoja
University of Turku, Turku, Finland
e-mail: hailme@utu.fi

B. Stocker, M. Mack (eds.), *The Palgrave Handbook of Philosophy and Literature*,
https://doi.org/10.1057/978-1-137-54794-1_28