

teorema

Vol. XXXV/3, 2016, pp. 185-199

ISSN: 0210-1602

[BIBLID 0210-1602 (2016) 35:3; pp. 185-199]

Fiction and Thought Experiment - A Case Study

Daniel Dohrn

RESUMEN

Muchos filósofos son muy optimistas respecto de la contribución cognitiva de la ficción a la ciencia y a la filosofía. Aquí, me concentro en un caso de estudio: la explicación que dan Ichikawa y Jarvis de los experimentos de pensamiento en términos de las historias de ficción ordinarias. En la medida en que la contribución de la ficción no es algo *sui generis*, el procesamiento de la ficción será a menudo parásito de las capacidades cognitivas que pueden reemplazarla; en la medida en que es *sui generis*, nada garantiza que la ficción se comporta suficientemente bien para acatar las constricciones del discurso del discurso científico y filosófico, por no hablar de los requisitos mínimos de coherencia conceptual y lógica.

PALABRAS CLAVE: *experimento de pensamiento, ficción, narrativa, necesidad, Gettier, contrafáctico, como si, fingir, historia.*

ABSTRACT

Many philosophers are very sanguine about the cognitive contributions of fiction to science and philosophy. I focus on a case study: Ichikawa and Jarvis's account of thought experiments in terms of everyday fictional stories. As far as the contribution of fiction is not *sui generis*, processing fiction often will be parasitic on cognitive capacities which may replace it; as far as it is *sui generis*, nothing guarantees that fiction is sufficiently well-behaved to abide by the constraints of scientific and philosophical discourse, not even by the minimum requirements of conceptual and logical coherence.

KEYWORDS: *Thought Experiment, Fiction, Narrative, Possibility, Necessity, Gettier, Counterfactual, As If, Pretense, Story.*

I. THE PURPORTED CONTRIBUTION OF FICTION

Fiction thrives in contemporary philosophy. Fictionalism looms large in discussing morals, modals, models, and so on. I distinguish two ways of using fiction. The first is cognitively highly regimented, subject to constant monitoring. In particular, it is constrained by the full apparatus of logics, mathematics, and scientific knowledge. One handy exam-

ple is fictionalism about mathematical objects, aiming only at minimising ontological commitments. One adds a fictionality operator, the intended effect being that one's commitments to numbers and so on do not 'count' in the same way as one's commitments to actual objects. It is an intriguing question how fictionalism relates to everyday fiction. Anyway this is not my topic in this paper.

I want to focus on the second use of fiction: use of ordinary stories in science and philosophy. By 'ordinary' I do not mean that the story must be commonplace but that there is no special philosophical regimentation as in fictionalism about some region of discourse. One takes some piece of scientific discourse and treats it as a story. There are pieces of scientific discourse which surely are amenable to being treated as stories, among them scientific thought experiments like Schrödinger's cat:

One can even set up quite burlesque 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 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), p. 157, translation slightly revised]

Though Schrödinger uses many scientific terms in a highly regimented way, his example would not have achieved notoriety without its 'burlesque' character: it can be read not only as a nomically possible scenario that exemplifies certain quantum phenomena but also as an entertaining science fiction short story. Though features like being burlesque or entertaining are not essential to fiction, everyday fictional stories are often crafted to display them.

There is a widespread tendency to think that thought experiments play an important role both in philosophy and science, and that this role is precisely due to features which they share with ordinary fictional stories, and which distinguish them from – in a sense – more rigorous forms of scientific discourse. Among these features which are typical for fictional stories (albeit not confined to story-telling) are pretense [Nichols and Stich (2000)], use of narrative patterns and certain modes of presen-

tation like metaphorical speech. Many philosophers credit these features with a special cognitive significance for scientific and philosophical inquiry [Camp (2009), p. 128], and especially for the scientific and philosophical practice of thought experimenting [Gendler (2007), p. 69]. Some even claim that (at least some) thought experiments (at least partly) are fictions, and that their cognitive significance is (at least partly) due to their being fictions [Carroll (2002), Meynell (2014)]. I shall consider one particular attempt to give everyday fictional stories a key role in processing a successful and significant philosophical thought experiment. I shall argue that this attempt is not successful, and that the problems it suffers from cast some general doubts on other attempts at using everyday fiction in scientific and philosophical inquiry.

II. AN EXEMPLUM CRUCIS

I choose an example of using fiction for the purposes of philosophical thought experiment from the literature. This example excels in being both more cautious and more precise than many of the usual appeals to everyday fiction. If the approach considered has difficulties, less elaborated ones are likely to have difficulties, too, but it will prove harder to bring out these difficulties clearly. I frame my considerations as an *explorative* meta-thought experiment: the general claim that fiction plays a genuine cognitive role is confronted with a test case.

My test case comes up in the context of a general metaphilosophical debate on how to formalise thought experiments like the Gettier experiment. Here is a paradigmatic Gettier scenario:

T1 At 8:28, somebody looked at a clock to see what time it was. The clock was broken; it had stopped exactly twenty-four hours previously. The subject believed, on the basis of the clock's reading, that it was 8:28. [Williamson (2009), p. 467]

Scenarios of this sort have convinced many philosophers that knowledge cannot be defined as justified true belief. In modal terms, it is not the case that knowledge is necessarily justified true belief. For the subject in the scenario intuitively has justified true belief but no knowledge (NKJTB) that it is 8:28.

Jonathan Ichikawa and Benjamin Jarvis propose to regard Gettier scenarios like T1 as *everyday literary fictions*. Let g be the proposition that

every element of the set of fictional truths in the T1-story is true (not just fictionally, true *tout court*). Then the Gettier argument can be resumed thus:

(1p) Possibly, *g*.

(2p) Necessarily, if *g*, then someone has NKJTB.

Therefore: (3p) Possibly, someone has NKJTB. [cf. Ichikawa and Jarvis (2009), p. 229]

Ichikawa and Jarvis see a strong connection between a suitable story and modal truth. But they are cautious. A set of truth-conditions is *extracted* from the story as corresponding to what is fictional. Then these truth-conditions are used to delimit a situation. The claim that this situation is possible must be established independently.

Ichikawa and Jarvis's use of fiction is designed to solve the problem of *deviant realisations*, which has been introduced by Timothy Williamson. Normal Gettier descriptions like T1 can be realised in a way which is deviant, i.e. does not serve as a counterexample as intended:

... T1 conjoined with the originally unintended T2 no longer works as a Gettier case: T2 The subject knew in advance that the clock had stopped exactly twenty-four hours previously. [Williamson (2009), p. 467]

So how are the *right* possible situations confined? Williamson himself famously proposed a counterfactual analysis. Roughly, if T1 had been actual, someone would have had NKJTB. However, it has been noted that the counterfactual analysis falls prey to the epistemic possibility that a deviant realisation is actual. Someone may actually instantiate T1 and T2. In that case the counterfactual might be false. But the thought experiment does not seem to depend on what is actually the case [Ichikawa (2009), Malmgren (2011)].

Ichikawa and Jarvis introduce fiction in order to select the non-deviant realisations of the Gettier scenario. In dealing with fictions we eschew deviant background conditions like T2. T2 is not true in a T1-fiction. Thus, Ichikawa and Jarvis say, our competence of evaluating everyday fictions allows us to single out the non-deviant possible situations the Gettier argument aims at. Ichikawa and Jarvis do not say much to motivate that the content of a Gettier fiction excludes deviant realisations. They rely on our competence to tell what is true in a fiction. For any par-

ticular deviant realisation one may come up with, we can tell that it is not true in the fiction. I am willing to grant this claim. If it is denied, we have even more reason to doubt the purported cognitive role of fiction.

Fiction is introduced into the thought experiment to deal with the problem of deviant realisations. Ichikawa and Jarvis are especially cautious in not letting fictional content but only the normal proposition g enter the Gettier argument. In the most cautious understanding of g , there is a one-to-one mapping from the set of fictional truths to some set of normal propositions. Proposition g says that the propositions in the latter set are true. We do not have to presuppose that the fictional truths are normal propositions.¹ This is an attractive feature for my case study; if the cautious approach leads to difficulties, any approach which directly inserts the content of the fiction into some philosophical argument is likely to lead to the same difficulties. For instance, it is argued that fiction serves to test our conceptual commitments [Carroll (2002)]. In order to do so, fiction must reliably preserve conceptual content. My case study shows that we cannot even rely on this minimum requirement.

One may question the significance of my case study. Ichikawa and Jarvis do not aim at establishing the general claim that fiction is cognitively useful in science and philosophy. Yet they are committed to the claim that using fiction is the best general way to deal with deviant realisations of thought experiments. Fixing thought experimental intuitions seems an especially important task. Moreover, we will see that the problems with the exemplary task generalise to other uses of fiction. My case study is not taken from the theory of fiction but from philosophical practice. In order to assess the general claim that fiction is cognitively useful, we should consider where fiction is actually used. It is highly significant to observe fiction actually being used for a well-defined theoretical purpose instead of merely speculating from a meta-perspective what fiction might be used for. I take Ichikawa and Jarvis's account to be a prime example of such a use (though I doubt their account of the role of fiction in thought experiments). My case study shares the limits of any case study. It might not be fully generalisable. If fiction fails with respect to the task considered, it may still be useful for others.

I shall consider two alternatives, depending on whether or not the contribution of everyday fiction to some cognitive endeavour, broadly understood, is *sui generis*. By *sui generis* I mean that processing fiction is more efficient in making this contribution than any alternative deployment of cognitive capacities. I do not claim that the contribution is due to essential features of fiction. Often the contribution of fiction is not *sui*

generis because it is parasitic on an ‘off-line’ use of cognitive capacities which are also used in dealing with the actual world. Then there is no reason why one should be expected to dwell on fiction rather than directly using the underlying cognitive capacities. If the contribution is *sui generis*, however, it will often be so because fiction is not regimented in the way required by philosophical and scientific inquiry. Then it is doubtful that the contribution will lead to the expected cognitive achievement. My first point (III.1) will elaborate on the first alternative, while my other two points (III.2-III.3) will elaborate on the second alternative.

III. CRITICISM

III.1 *Fiction as an Epistemic Detour*

My first criticism targets the special epistemic significance of fiction. Fiction is often vaguely credited with promoting ‘understanding’ or ‘explanation’ [Camp (2009), pp. 127-28]. I do not deny that it may contribute in its own way to cognitive achievements. Yet the question one would have to address is whether there is anything about a cognitive endeavour that makes fiction especially suited to promote it, compared to a more direct approach to real or possible scenarios. Use of epistemic capacities ‘off-line’, say within a game of pretense, will often be parasitic on their ‘on-line’ use in dealing with real scenarios. Fiction is dispensable as far as one might as well directly use the capacities which underlie our dealing with fiction. When fiction is dispensable, we may doubt that it is used at all. The general suspicion that fiction is an idle wheel is illustrated by my case study. It could be dispelled by presenting a genuine function of fiction in philosophical inquiry which is not parasitic on the use of cognitive capacities outside of fiction. But here the burden of proof is on the part of the advocate of fiction.

An exemplary target which makes this principled criticism more concrete is Ichikawa and Jarvis’s claim that in dealing with fictions, we discard deviant background conditions. If we are able to know (1p) and (2p), the premisses of the Gettier argument à la Ichikawa and Jarvis, we must at least in principle be able to disentangle deviant and non-deviant background conditions. The problem is not that we cannot refer to the right situations by using *g*, the proposition fixed by the fiction, but that we cannot know that these situations are NKJTB-situations without having an independent access to them. We must have a suitable grip on what

all the T1-scenarios under discussion are like to see that they are all NKJTB-scenarios. One may deny that this issue is to be settled by the capacity of evaluating fiction. But whatever additional capacity takes over, it must allow to consider precisely the scenarios selected by the fiction. I do not see how this condition could be guaranteed to be met unless evaluating fiction piggybacks on an independent cognitive capacity of considering precisely the right situations. Frank Jackson hints at such a capacity when he compares thought experiments to cooking recipes:

we make an assumption akin to that we make when we follow a cooking recipe. Cooking recipes don't tell you what *not* to do. They couldn't. There are infinitely many things not to do (don't add bats wings, don't arrange for the power to be turned off, don't burn the house down, ...). They tell you what to do, and rely on your having an implicit grasp of an 'and that's all' clause. The same goes for described possible cases. They are given in a relatively small number of words. This of necessity leaves many things open. Those presenting the cases rely on hearers having an implicit grasp of the range of cases being presented [Jackson (2011), p. 478]

Assume for the sake of argument that a 'that's all'-clause is all that it takes to get rid of deviant realisations. The corresponding feature of fiction is the incompleteness of fictional stories. For instance, there is no number n such that it is true in Nabokov's *Lolita* that n people were born in Paris in 1910 [cf. Woodward (2012)]. Analogously, the T1-story may be incomplete with respect to T2 and other deviant realisations. It is neither true that, in the T1-fiction, T2, nor is it true that, in the T1-fiction, not T2. In contrast, possible worlds are taken to be complete: for any proposition p and any world w , either p is true or it is false at w . Hence, if we just take the worlds where what is true in the T1-fiction is true, we do not thereby exclude the T2-worlds. Ichikawa and Jarvis would need a recipe what to do with things that are indeterminate in the fiction. Thus, a 'that's all'-clause won't do.

I do not think anyway that a 'that's all'-clause could take care of deviant ways of fleshing out a scenario. For there are things one must *do* without being explicitly told so: before cooking, ensure that the power is turned on. In a similar vein, there are untold things we must assume for the T1-scenario to work: clocks are good indicators of time. Jackson's analogy to the cooking recipe raises the question of how we tell apart deviant and normal ways of following the recipe. A proposal analogous to Ichikawa and Jarvis's theory would be that we dwell on a literary fiction of cooking to figure out how to implement the recipe. But this

seems preposterous. We usually do not need a fiction to fix the content of a cooking recipe (i.e. what is and what is not to be done according to the recipe). We rely on an implicit sense of normalcy. In processing fiction, we also deploy a normalcy heuristic, but deploying a normalcy heuristic does not depend on fiction. The delicate and ill-understood capacity of telling apart normal from deviant circumstances plays a great role in everyday life. And invoking fiction in order to use this capacity seems to be an unnecessary detour. There are other ways of explicitly imposing normalcy conditions, for instance ‘*ceteris paribus*’, ‘normally’. The Gettier argument becomes: let g' be the proposition that anything that is *normally* the case in a T1-scenario is true.

(1n) Possibly, g' .

(2n) Necessarily, if g' , then someone has NKJTB.

Therefore: (3n) Possibly, someone has NKJTB.

The normalcy clause performs the task the fiction is supposed to perform.²

To summarise: in order to bring out the special epistemic role of fiction, we should ask in how far its exploits are *genuine* or parasitic on more mundane epistemic capacities.

III.2 Narrative Persuasion vs. Argument

A standard objection against the cognitive role of fiction is *No Argument*: neither artworks nor discourse about them are bound to provide arguments for claims about the world [Carroll (2002), p. 6, Lamarque and Olsen (1994), pp. 332-3]. *No Argument* is only half of the truth. The other half is that fiction is highly sensitive to certain mechanisms of persuasion which do not abide by scientific standards. According to recent research on *narrative transportation*, emotions of empathy and imagery elicited by a story dispose the audience to get detached from actuality and immersed into fictive scenarios [Van Laer et al. (2014), Liao and Doggett (2014)]. One's engagement with fiction tends to have a persuasive impact on one's view of the actual world, even when the story itself is not overtly persuasive. This persuasive role of narratives often does not conform to scientific and philosophical standards of reflectively controlled methodical reasoning. I do not dismiss narrative transportation as irrational. It

simply is not in the business of providing well-regimented philosophical arguments. David Velleman introduces a corresponding notion of narrative (pseudo-)explanation ('Aha! Of course!', [(2003), p. 21]). As an example, he presents a story where an usurper is crushed by a statue of his predecessor. We tend to perceive the usurper's death as a punishment caused by his misdeed. Of course, this may be explained by conventions of poetic justice, but one may be subliminally tempted to project corresponding pseudo-explanatory relationships into similar actual events. Narrative explanation is not confined to fiction, but everyday fictional stories are especially amenable to it.

Velleman's example can be used to illustrate the problem for treating philosophical scenarios as fictions. Given normal conventions of story-telling, it is not absurd to read the pseudo-explanation into the usurper story. In this reading, it is true in the story that the usurper was punished according to some natural mechanism of poetic justice. But if we were to read the story as a philosophical thought experiment, we should be more careful. For instance, there are doubts whether such a mechanism is metaphysically possible. The usurper story is no philosophical thought experiment. Still it illustrates how narrative persuasion may lead to unexpected differences between a scenario treated as a philosophical thought experiment and as an everyday fiction. Due to these differences, one cannot simply take for granted that treating a scenario as a fiction will yield the set of possibilities intended by the thought experimenter. The fission scenario to be considered in the next section will further illustrate the point.

III.3 *Fictional Content Is Not Guaranteed to Be Well-Behaved*

The argument from narrative explanation gave rise to the suspicion that the fiction carves out the wrong set of possibilities instead of those relevant to a corresponding thought experiment. The argument to come leads to a similar conclusion by more generally addressing the restrictions imposed on fictional content by its purported use in philosophy. Ichikawa and Jarvis's account presupposes that there is a precise correspondence between the set of fictional truths and the set of possible situations relevant to a corresponding thought experiment. This presupposition is not supported by the extant accounts of truth in fiction mentioned by Ichikawa and Jarvis [Ichikawa and Jarvis (2013), p. 265]. In Lewis's (1983) counterfactual analysis, roughly p is true in a fiction iff it is true in the closest world where the story is told as known fact (variant: ... and

the common beliefs of the community of origin are true). However, Ichikawa and Jarvis motivate their fictional account precisely by the failure of Williamson's counterfactual formalisation of thought experiments to deal with the possibility of an actually deviant realisation. If Williamson has problems with letting actually deviant realisations in, a fictional account drawing on Lewis's analysis would have problems, too. Moreover, the counterfactual account of fiction threatens the fictional account with redundancy. One should expect that an outright counterfactual can replace the fiction in most philosophically relevant contexts.

There are alternatives to Lewis's analysis. In Kendall Walton's (1990) pluralistic account, the explicit story functions as a *prop*. It invites us to engage in a game of *make believe*. One is authorised to make believe precisely the fictional truths, which are determined according to several conventional principles of generation. As Walton himself emphasises, truth in fiction is not a kind of truth [cf. Walton (1990), p. 42]. Walton's principles of generation are not bound to ensure that the content of a *make believe* game can be mapped to propositional truths. Hence philosophers who want fiction to provide the content of a thought experiment cannot use Lewis's or Walton's accounts but must come up with a picture of their own.³ The picture used by Ichikawa and Jarvis is rather minimalistic. I shall show that even this minimalistic picture leads to grievous doubts about the cognitive use of fiction.

How do Ichikawa and Jarvis ensure that the set of fictional truths corresponds to the right set of possibilities? A key (but not sufficient) requirement is the following: for any coherent thought experimental scenario, the set of truths which correspond to what is true in the description of the scenario, treated as an everyday fiction, must also be perfectly coherent, interpreted broadly as *logical consistency and conceptual coherence*. In other words, truth in fiction is subject to the same constraints of preserving conceptual and logical truths as modal reasoning [cf. Ichikawa and Jarvis (2009), p. 234; p. 237]. If we accept the idea of conceptual coherence, this minimum requirement seems crucial to any account that somehow uses fiction to fix the content of a thought experiment. For assume there is a thought experiment where the coherence constraint is not met. Some description is perfectly coherent when treated as a thought experiment but not as a fictional story. Then we would clearly be bound to settle for the coherent scenario in the philosophical argument. But Ichikawa and Jarvis's procedure would lead us to diagnose an incoherence: the thought experiment fails because there is no coherent

proposition according to which what is true in the fiction is true *tout court*. This diagnosis would obviously be misplaced.

Disregarding concerns about its sufficiency, I shall argue that even the minimum requirement of conceptual coherence is not always met. Constraints of conceptual and logical coherence are weaker and thus more easily overridden by other criteria in everyday fiction than in thought experiments. Fiction may not be arbitrarily incoherent, but it tolerates suitably embedded punctual incoherence. As Gregory Curry notes, one may write a fiction where it is explicitly told that someone has refuted Gödel's theorem [cf. Currie (1990), p. 69]. So far there is no problem for the fictional account. The fiction is inconsistent, and so would be a corresponding thought experiment. The problem arises when the story is *implicitly* incoherent.

One may write a fiction where it is not explicit (or entailed) but *only implicitly* true that someone has refuted Gödel's theorem. Assume the author of the story wants to make vivid what a superb genius her protagonist Schmidt is. She elaborates the ceremony where Schmidt is awarded the Fields medal for having refuted Gödel's theorem.⁴ If the story is suitably told, no hoax, irony, or indication of error, we are quite ready to accept it as true according to the story that Schmidt *has* refuted Gödel's theorem. Narrative plausibility (Velleman's 'aha; of course!') trumps conceptual coherence and logical consistency. But the same would not follow in a literally identical thought experimental scenario where metaphysical possibility and *a fortiori* conceptual coherence are strictly to be preserved. One way of securing these constraints in the Smith story would be to interpret the story such that the laureate and the committee *made a mistake*. We are bound to an interpretation like this if coherence is non-negotiable but not in dealing with fiction.⁵

In sum, the purported correspondence between fiction and thought experiments fails. What is perfectly coherent as a thought experiment may become incoherent as an everyday fiction.⁶ Although my argument is sufficient to shed doubt on the fictional account, there is a fallback position: one may claim that the requirement of conceptual coherence is fulfilled for any interesting philosophical thought experiment. For any such thought experiment, the corresponding everyday fiction is coherent. One concern is that it seems irresponsible to use fiction as long as there is no positive argument that this fallback position is correct. Instead of further pressing this concern, I shall test the fallback position by going through a concrete counterexample. My example will be *fission cases*. I shall evince how the fictional account may misconstrue the dialectics of

such cases. The outline of a fission case is the following: imagine a person P0, whose brain is divided into two halves and implanted into two bodies; two normally functioning persons emerge, P1 and P2, who are psychologically continuous with P0. Consider the following story template (I leave out the proper names to make a purely non-fictional reading available):

T3 Someone undergoes fission. Both of the post-fission persons sincerely utter: ‘I remember the slightest details about *my* pre-fission life. I remember my early childhood, my grandparents when they were still alive (...story continues).

In that story, drawing on psychological continuity, both protagonists P1 and P2 seem to refer to P0 as ‘I’. It is a highly intricate and disputed question who, if anyone, is right. Before the dispute is settled, any participant in the debate should read the scenario such as to preserve a coherent common ground: P1 and P2 *take* themselves to be identical to P0. Any stronger claim to identity is suspicious of violating conceptual boundaries [Sattig (2012)]. If conceptual coherence is to be strictly preserved, one should eschew such a stronger claim as long as its conceptual coherence is doubtful.

However, when we read T3 as a science fiction story, we do not seem bound by these philosophical considerations. If the self-image of the protagonists is sufficiently vivid, we may accept it as fictionally true that P1 was P0 and/or that P2 was P0. Still the story does not explicitly say so. Given these assumptions, we get a counterexample to the correspondence thesis. Read within the constraints imposed on philosophical reasoning, T3 describes a perfectly coherent and possible scenario. But our reading of T3 as a science fiction story is not bound to be perfectly coherent. The reason is not that fiction is not principally bound by conceptual coherence. It is bound to be coherent, other things being equal. But in the fission case, other things are not equal. The relevant conceptual constraints are so subtle that only a thorough philosophical inquiry of fission cases and the like discloses them. One cannot be expected to observe them in competently evaluating a normal science fiction story. Science fiction may invite highly sophisticated conceptual distinctions, but that would require more subtle clues than a simple fission scenario as used in philosophy. One may wonder how one can even grasp the story if it is not in tune with our notion of personal identity. One way to account for our grasp would be to draw on an initial understanding which

does not require a full mastery of any subtle conceptual truth about personal identity.

However, shouldn't the author's intention of presenting a coherent story constrain our interpretation of *the fiction* such as to prevent the incoherent reading? No, that intention is irretrievably bound to the aim of presenting a philosophical argument. But that aim is blinded out when the story is treated as an *everyday* fiction.

One last fallback position would be the following: only successful thought experiments show a perfect correspondence with a literally identical story, and the success of the fission story is contestable. But given the difficulty of singling out a successful thought experiment, this fallback option smells of an immunisation strategy.

I conclude that the fictional account is flawed. Fiction is not well-regulated enough to be used in a general analysis of Gettier-like thought experiments. In how far are these results paradigmatic? Firstly, our competence of closing in on the right fictional scenarios is parasitic on a competence of imposing normalcy conditions. Generalising, any candidate for a specific cognitive achievement of fiction is suspicious of being an epistemic detour because it is parasitic on other epistemic capacities which are of use outside of fiction and can as well be independently activated without invoking fiction. Secondly, fiction is not well-behaved philosophically. It makes things narratively plausible which are not plausible in the sober environment of a philosophical argument. Not even a scenario which would be perfectly coherent as a philosophical thought experiment has to be logically and conceptually coherent, taken as an everyday fiction. But if narrative plausibility may trump even a minimum condition like logical and conceptual coherence, use of ordinary fiction in philosophical reasoning seems deeply problematic.

I close with a note in caution: none of my results preclude that we may come up with highly regulated philosophical pieces of fiction, as in the many variants of philosophical fictionalism. However, even concerning this highly regulated use, there is a suspicion that, to the extent that it simply relies on our everyday capacity of dealing with ordinary fictions, it may be infected by the latter's irregularities.

*Institut für Philosophie
Humboldt-Universität zu Berlin
Unter den Linden 6, 10099 Berlin, Germany
E-mail: daniel.dohrn@hu-berlin.de*

NOTES

¹ Ichikawa and Jarvis consider using fiction as a reference fixer for a demonstrative ‘things are like that...’. One may doubt that they are committed to a mapping from fictional truths to truths. But I reckon such a mapping the only way of ensuring that *g* selects the right possibilities.

² Malmgren criticises that such clauses cannot be cashed out [Malmgren (2011), pp. 286-7]. But it is not clear that we must be able to cash them out as long as we are able to recognise deviant completions of a recipe when we encounter them. Be that as it may, I do not claim that a normalcy clause solves the problem, just that it does as well as a fiction. If the normalcy clause cannot be cashed out, the same goes for truth in fiction.

³ I do not reckon more recent accounts of fictional truth fit for the task. Nichols and Stich (2000), understood most charitably, construe fictional truth by a process of revising one’s actual belief system by the explicit story. But since it does not follow from our beliefs that T1 is not deviantly realised, e.g. as T2, this won’t help to close in on non-deviant scenarios. Other accounts do not seem promising either [overview in Alward (2009)]. For instance, they often come up with an ideal author or audience. But ‘ideal’ is not understood as ‘ideal for the purposes of a philosophical argument’.

⁴ It is important that ‘for ...’ is read intensionally: the prize committee’s opinion that Schmidt has refuted Gödel’s theorem is their reason why they award him the medal. For related examples [van Inwagen (1998), p. 79, Kung (2016), p. 96].

⁵ I have been reminded that for instance in an ethical trolley experiment where the issue is whether to save Schmidt, who actually refuted Gödel’s Theorem, or three other people, we might prefer to accept that Schmidt has refuted Gödel’s Theorem. But this is only because we see that this feature is irrelevant to the story. My suspicion is that we practice some sort of double-accounting. We officially stick to the story, but we ignore the weird but transparently idle detail when it comes to evaluating the argument.

⁶ I have encountered doubts about the idea of a ‘literally identical’ description, read as a thought experiment and as a fictional story. But the intuitive idea is plausible: T1 can of course be read as part of an epistemological argument. And our conventions of telling fictional stories seem flexible enough to embed T1 into a speech-act of story-telling (which may be represented by adding an ‘according to the fiction’-operator).

REFERENCES

- ALWARD, P. (2010), ‘That’s the Fictional Truth, Ruth’, *Acta Analytica*, vol. 25, pp. 347-63.

- CAMP, E. (2009), 'Two Varieties of Literary Imagination', *Midwest Studies in Philosophy*, vol. 33, pp. 107-30.
- CARROLL, N. (2002), 'The Wheel of Virtue: Art, Literature, and Moral Knowledge', *Journal of Aesthetics and Art Criticism*, vol. 60, pp. 3-23.
- GENDLER, T. (2007), 'Philosophical Thought Experiment, Intuitions, and Cognitive Equilibrium', *Midwest Studies in Philosophy*, vol. 31, pp. 68-89.
- ICHIKAWA, J. (2009), 'Knowing the Intuition and Knowing the Counterfactual', *Philosophical Studies*, vol. 145, pp. 435-43.
- ICHIKAWA, J. and JARVIS, B. (2009), 'Thought-experiment Intuitions and Truth in Fiction', *Philosophical Studies*, vol. 142, pp. 221-46.
- (2013), *The Rules of Thought*, Oxford, OUP.
- INWAGEN, VAN, P. (1998), 'Modal Epistemology', *Philosophical Studies*, vol. 92, pp. 67- 84.
- JACKSON, F. (2011), 'On Gettier Holdouts', *Mind and Language*, vol. 26, pp. 468-81.
- KUNG, P. (2016), 'You Really Do Imagine It. Against Error Theories of the Imagination', *Noûs*, vol. 50, pp. 90-120.
- LAER, VAN, T. et al. (2014), 'The Extended Transportation-Imagery Model: A Meta-Analysis of the Antecedents of Consumers' Narrative Transportation', *Journal of Consumer Research*, vol. 40, pp. 335-43.
- LIAO, S. and DOGGETT, T. (2014), 'The Imagination Box', *Journal of Philosophy*, vol. 111, pp. 259-75.
- LEWIS, D. (1983), 'Truth in Fiction', in *Philosophical Papers I*, Oxford, OUP, pp. 261-80.
- (1986), *On the Plurality of Worlds*, Oxford, Blackwell.
- MALMGREN, A. S. (2011), 'Rationalism and the Content of Intuitive Judgements', *Mind*, vol. 120, pp. 263-327.
- MEYNELL, L. (2014), 'Imagination and Insight: A New Account of the Content of Thought Experiments', *Synthese*, vol. 191, pp. 4149-4168.
- NICHOLS, S. and STICH, S. (2000), 'A Cognitive Theory of Pretense', *Cognition*, vol. 14, pp. 115-47.
- SATTIG, T. (2012), 'The Paradox of Fission and the Ontology of Ordinary Objects', *Philosophy and Phenomenological Research*, vol. 85, pp. 594-623.
- SCHRÖDINGER, E. (1935 [1983]), 'The Present Situation in Quantum Mechanics', in Wheeler, J.A. and Zurek, W.H. (eds.), *Quantum Theory and Measurement*, Princeton, Princeton University Press, pp. 152-68.
- VELLEMAN, D. (2003), 'Narrative Explanation', *The Philosophical Review*, vol. 112, pp. 1- 25.
- WALTON, K. (1990), *Mimesis as Make-Believe. On the Foundations of the Representational Arts*, Cambridge (Mass.), Harvard University Press.
- WILLIAMSON, T. (2007), *The Philosophy of Philosophy*, Oxford, Blackwell.
- (2009), 'Replies to Ichikawa, Martin and Weinberg', *Philosophical Studies*, vol. 145, pp. 465-76.
- WOODWARD, R. (2012), 'Fictionalism and Incompleteness', *Noûs*, vol. 46, pp. 781-90.