Knowledge from fiction and the challenge from luck

Article (Accepted Version)

Stock, Kathleen (2019) Knowledge from fiction and the challenge from luck. Grazer Philosophische Studien, 96 (3). pp. 476-496. ISSN 0165-9227

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/84952/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Knowledge from fiction and the challenge from luck

Abstract

In order for true beliefs acquired from reading fiction to count as knowledge proper, they must survive 'the challenge from luck'. That is, it must be established that such beliefs are neither luckily true, nor luckily believed by readers. I consider three kinds of true belief a reader may, I assume, get from reading fiction: a) those based on testimony about empirical facts; b) those based on 'true in passing' sentences; and c) those beliefs about counterfactuals one may get from reading a 'didactic' fiction. The first group escape the challenge from luck relatively easily, I argue. However, things turn out to be more complicated with the second group. I examine Mitchell Green's suggestion, effectively, that knowledge of *fictional genre* may see off the challenge from luck here, but reject this in the form presented by Green, adapting it substantially to offer beliefs of this kind a more promising escape route. I finish by following Green's lead once again, and discussing the category of 'didactic' fiction, as he calls it. I argue that any true beliefs about counterfactuals gained from such fictions are likely to be lucky. I conclude however that things are much more promising for any true beliefs gained about oneself as a result of engaging with what Green calls an 'interrogative' fiction.

Keywords

Fiction; knowledge; cognitivism; Mitchell Green; luck epistemology; epistemology of fiction.

Author bio:

Kathleen Stock is a Professor in Philosophy at the University of Sussex, UK. She has published widely on the philosophy of fiction and imagination. Her monograph *Only Imagine: Fiction, Interpretation and Imagination* (2017) is out now with Oxford University Press.

Knowledge from fiction and the challenge from luck

1 Literary cognitivism

Literary cognitivists claim that readers can get interesting knowledge about the world from fictions: from novels, short stories, TV dramas, and so on. For instance, a reader can apparently get new knowledge of empirical facts. Here's an illustrative passage from Karen Joy Fowler's novel *We Are All Completely Beside Ourselves*. It informs the reader about some concrete historical events concerning the Animal Liberation Front, which presumably many did not know before:

Nonhuman animals have gone to court before. Arguably, the first ALF action in the United States was the release of two dolphins in 1977 from the University of Hawaii. The men responsible were charged with grand theft. (Fowler 2013, 305)

A reader also might seem to get knowledge from fiction, even where the information in question isn't introduced quite as explicitly. Reading Laura Ingalls Wilder's *By The Shores of Silver Lake* as a child, I learnt some of the ingredients for sourdough biscuits. From Louisa May Alcott's *Little Women* I learnt that in the 19th Century, there was a market for selling one's hair. From Dodie Smith's *I Capture the Castle* I learnt that in the Old Testament, Jacob wrestles an angel; and that crème de menthe is a green alcoholic liqueur. From Agatha Christie books, I learnt that curare is a poison that can be placed on an arrow tip. Presumably there are many other facts I now 'know' whose fictional origins I have forgotten. Unlike in the Fowler case, none of these are introduced, within their associated works, in a guise that particularly announces them as 'facts to be believed', however. They are put into the mouths of characters, or fallible narrators; they're implied rather than stated; they're mentioned 'in passing' as illustrative of other points. I too picked them up in passing, as it were.

A third sort of case seemingly friendly to literary cognitivism is that of fiction which furnishes knowledge about counterfactual truths. In a series of stimulating papers (2008, 2010, 2016, 2017), Mitchell Green closely explores this aspect of fiction. Green defines literary cognitivism more narrowly than I just did, as the claim that 'literary fiction can be a source of knowledge *in a way that depends crucially on its being fictional*' (2010, 352; my italics). Any knowledge of counterfactuals derived from fiction would seem to depend 'crucially' on the fictionality of the vehicle, and so satisfy literary cognitivism as Green more strictly characterises it.

Green argues that, in effect, we should treat a significant number of fictions as thought experiments. Each presents a hypothetical situation, to 'determine what consequences may plausibly follow' (2010, 360). Moreover, amongst these fictions, we should distinguish between the 'interrogative' ones – fictions which ask questions about what would be the case, given a certain

starting point, thereby prompting the reader to her own reflection, without some specified conclusion being implied – and those which are ‘didactic’ (Green 2016, 289).

‘Didactic’ fictions positively imply some determinate counterfactual: if *that* happened, then *this* likely would, or at least could, happen. So for instance, Aldous Huxley’s *Brave New World* effectively says that ‘were society to exist based on principles associated with hedonic utilitarianism’ then ‘the State would support an orgiastic religion, drug-induced anaesthetization of its citizens, suppress all independent inquiry, and so forth’ (Green 2010, 360). And Ursula Le Guin’s *The Left Hand of Darkness* concretely describes the possible consequences for human relationships of a scenario in which people change gender on a regular basis (Green 2017, 54). Counterfactuals implied didactically by fictions can be true, Green argues, and some are: they describe real possibilities or probabilities, given certain antecedents. Readers who believe the counterfactual implications of such a narrative get knowledge. Or so he claims.

2 The challenge to literary cognitivism from luck

All such claims face a distinctive challenge, which I’ll spend the rest of this essay exploring. This challenge is based on the assumption that, for some proposition ‘*p*’, knowing that *p* requires, not just that one believe that *p* and that one’s belief be true, but also that one’s belief cannot be *luckily* true.

Epistemologists tend to characterise a luckily true belief in terms of it being a live possibility – i.e. reasonably likely - that the believer could have formed the very same proposition as a belief, under relevantly similar conditions, and yet that belief be false. See, for instance, Duncan Pritchard:

S’s true belief is lucky iff there is a wide class of near-by possible worlds in which S continues to believe the target proposition, and the relevant initial conditions for the formation of that belief are the same as in the actual world, and yet the belief is false (2007, 280)

An oft-recounted and vivid imaginary example of a lucky true belief, which satisfies this description, is that of a person wandering about in Fake Barn Country. In this country, as the name suggests, there are many fake barn facades indistinguishable from real ones. One day, the wanderer comes across a real barn, and believes, correctly, that she does so. Nonetheless, it is worried, she does not *know* that there is a real barn in front of her, because there are close possible worlds in which she is standing in front of an identical fake barn, yet has a belief with identical content. That is, it seems, her true belief is lucky. Other standard examples include lucky wild guesses about fairly improbable events; accidentally true beliefs based on wishful thinking rather than evidence or reasoning; and beliefs based on hearing lies which are, unbeknownst to their utterer, true.

Let’s assume we do gain true beliefs from fiction, both actual and counterfactual. I won’t question this in what follows. However, we might still ask: are the true

beliefs we gain from fiction about empirical facts lucky in the sense just given, or any near relation?¹ If so then, according to most, those beliefs couldn't count as knowledge proper, even though true. It seems that there are two broad reasons for which true beliefs acquired from fiction might look worryingly lucky, at least on the face of it.

The first involves a worry that true beliefs from fiction are lucky in exactly the sense just described by Pritchard. Two distinct scenarios can be cited in evidence. One of these involves doubt about the reliability of authors with respect to the facts they are writing about². If authors of fictions tended to be particularly unreliable about a given domain of facts, then it would seem to make whatever true sentences they stated or implied about that domain in their fictions look accidentally true. That is, there would be close possible worlds in which they produced the same sentences, but those sentences were false. If a reader then believed those sentences, then her beliefs would be only luckily true.

Another scenario in which true beliefs acquired from fiction look lucky involves the background assumption that the practice of creating fiction is incompatible with one's ultimate goal being to tell the truth. For instance, Harry Deutsch argues that fiction, as such, is not primarily intended to track or reflect what is actually true; even though he happily allows that, as a matter of fact and for non-truth-related reasons, a fiction may contain 'generous amounts of .. fact woven in as well' (2000, 167). Equally, David Davies has stated that:

'..whether or not the narrated events are true or known to be true, their having occurred is not relevant to what the author [of a fiction] is trying to achieve in writing the narrative... The proposition's being true is not the reason for its inclusion in the narrative'. (Davies 2007, 47)³

If this is right, then wherever an author of a fiction tells or implies the truth in her work, she does not do so, ultimately, for truth-telling reasons, but rather in the services of some other goal. It would seem to follow perhaps that whatever truths a fiction contains might easily have been falsehoods, as long as they still served that further non-truth-related goal⁴. Here too then, it might seem, a reader gains a true belief that *p* as a result of reading fiction, yet there are close possible worlds where she believes that *p*, via relevantly similar routes, and her belief that *p* is false.

Perhaps a reader's true beliefs, acquired from fiction, are lucky in a different and extended sense. This time we aren't so much concerned with what the author

¹ Stacie Friend pursues though ultimately does not endorse a related line, in her 2014. I respond in my 2017a.

² I don't mean to imply that a speaker's being reliable is a necessary condition of their transmitting knowledge via testimony; but only that unreliability is destructive of transmission (Adler 2017 Section 3).

³ This isn't a view he now holds.

⁴ The wider question of the relation between fiction and truth is discussed in Stock 2017b, Chapter 5.

might easily have done, but with what the reader might easily have done. In this scenario, readers are only accidentally able to identify likely-to-be-true sentences or implications within fictions. That is to say, there are close possible worlds in which the reader, instead of identifying as likely-to-be-true, and so believing, a true sentence or implication that *p*, present in a fiction, identifies as likely-to-be-true and believes one of the false sentences surrounding it instead. In these worlds, there's not a belief with precisely the same content *p*, but there's still a false belief instead of the true belief that *p*, arrived at by relevantly similar routes. We might use an analogy: if a person has true and false sentences randomly fired at her, concerning a factual domain she knows nothing about, and she somehow believes one of the true ones, this looks lucky, insofar as she might easily have believed one of the false ones.

Support for this scenario might seem to come, fairly basically, from the sheer number of false sentences normally present in a fiction. Given this proliferation, it might seem that the reader must be unlikely to be able to identify the likely-to-be-true ones reliably. We might also add the point made by Deutsch and Davies just now. For if, as was claimed, the practice of writing fiction is incompatible with truth-telling as a ultimate goal, then it looks like the reader now has no access to one familiar reason for thinking that certain true sentences are likely-to-be-true. Namely, generally speaking, once a person has identified an utterance as motivated by the intention to tell the truth, this often gives her at least some reason to think it true, for:

‘... the mere fact that someone would be willing to go to the trouble of asserting *P* at all is a *prima facie* reason for thinking that *P* is true.’
(Pritchard 2004, 332)

Without this, then, perhaps the reader is only accidentally hitting the mark when she believes something true as a result of reading fiction.

3 Testimony-in-fiction and escaping the challenge from luck

In the previous section, I used an example from the novel *We are All Completely Beside Ourselves*, as follows:

Nonhuman animals have gone to court before. Arguably, the first ALF action in the United States was the release of two dolphins in 1977 from the University of Hawaii. The men responsible were charged with grand theft. Their original defense, that dolphins are persons (humans in dolphin suits, one defendant said), was quickly thrown out by the judge. (Fowler 2013, 305)

This, I suggest, is an example of ‘testimony-in-fiction’. As such, I will argue, it escapes the challenge from luck described above. There is no reason to think that the author is unreliable about the facts she relates here, and in fact, as we’ll see, there is some reason to think her particularly credible. There is good reason to think that her goal in this passage is ultimately to tell the truth. And there is no reason to think that the competent reader cannot reliably identify these

sentences as likely-to-be-true, and differentiate them from the surrounding falsehoods in the text. This example is not unusual in these respects; it is fairly typical of testimony-in-fiction. Hence cases of testimony-in-fiction normally escape the challenge from luck.

First, a brief word on testimony, generally. Most testimony takes place in non-fictional contexts of utterance. Though philosophical accounts of testimony are contested, we can at least say that generally speaking, testimony that 'p' necessarily involves the utterer of 'p' i) intending to induce a belief that *p* in the hearer (or reader) of her utterances; and ii) intending that her utterance 'p' be a source of independent warrant for the hearer's (or reader's) belief in the truth of its content. In normal and well-formed cases, moreover, testimony can be a genuine way of transferring knowledge about a factual matter, without the hearer having to additionally check that what she hears is true. (Controversy exists about the background conditions required for well-formed cases; see Adler 2017 for an overview).

I'll now explain how the Fowler example withstands the challenges listed above. The first one was a general worry that authors might be unreliable about the matters they describe, so that even if they sometimes hit the mark and say something true, they might easily have said something false. In the case of the example we're discussing, however, there's no particular reason to think that Fowler is unreliable about the matters she relates, and indeed, there is reason to think she's reliable. Her acknowledgements cite academics from the 'Chimpanzee and Human Communication Institute in Ellensburg, Washington' and 'Dr Carla Frecerro for her readings and lectures on animal theory' (Fowler 2013, 310). In an epilogue, she writes:

Initially I was very nervous about all the things I didn't know regarding chimpanzees. Fortunately, there are any number of books written by the people involved in these studies. I did my best to read them all (Fowler 2013, 314).

In other words, she has apparently researched her subject deeply and with care. It would be strange if this passage was an exception.

A second part of the challenge from luck was the worry that, where authors of fictions include true sentences in their works, they don't do so for truth-telling reasons, ultimately. However, this too seems false for our example. For Fowler, it is clearly central to her goal in writing this passage to tell the truth by means of it. She isn't telling the truth for some goal which might have been satisfied just as well by a falsehood. Since part of her aim in the novel is to get readers to reconceptualise or at least complicate their view of actual 'real-life' higher mammals and their relation to personhood, it clearly complements this general aim to convey to the reader true historical information about relevant cases. A mark of this is that the reader, reading this passage in its original context, would feel entitled to criticise Fowler, should it emerge that she had got the facts wrong.

Finally, and relatedly, there's no reason to think that the reader can't non-accidentally distinguish Fowler's sentences as likely-to-be-true, and differentiate them from surrounding false ones⁵. She can do this by non-accidentally noticing Fowler's intention to inform by means of this passage. In earlier work, I have suggested that, in cases like this, the competent reader can fairly easily discern sentences that she is intended by the author to believe (Stock 2017a, 29). Though I don't pretend to give a complete account of the means by which this occurs, they include – as in the Fowler case – the inclusion of sentences which satisfy some or all of the following: they are in the declarative mood, or imply a declarative statement; they possess an authoritative-sounding tone; they are concerned with real, actual existents, recognisable as such; they complement and extend information the reader already possesses, and are recognisable as such; they contain information that is of use, interest, or relevance to the reader, and are recognisable as such; and they are not accompanied by identifiably invented content about the very same entities (*ibid.*). Against a shared understanding of such methods, Fowler is able to intentionally communicate to the reader, on the one hand, what she wants her to believe; and on the other, what she wants her only to imagine, and not believe.

With true testimony-in-fiction, then, as with this case, there's normally no easily conceivable close possible world in which the author might have said something false by the same utterance and yet the reader believed it. Equally, the competent reader is normally able to reliably and non-accidentally pick out the sentences that are likely-to-be-true from those that aren't. Her resulting true beliefs therefore don't look lucky, and may count as knowledge.

4 'True-in-passing' sentences and the challenge reasserted

This still leaves a large swathe of other true beliefs gained from fiction unaccounted for. Earlier, I noted how, in reading as a child, I apparently learnt from novels facts such as those concerning the ingredients of sourdough biscuits, and the fact that in the 19th Century, there was a market for selling one's hair. Arguably, though, and unlike in the cases discussed above, the authors of the works concerned didn't specifically intend to teach readers about these points. We can tell this partly because such facts, where they're introduced, do not possess many or even any of the signs described above as indicative of the intention to inform and to be recognised as such. They are not presented in the declarative mood in an authoritative way, but only in passing, often by a character or manifestly fallible narrator; they don't particularly complement the theme of the book in which they feature; they don't contain information reasonably presumed as of interest to the reader; and arguably there would be no particular cost to the reader, should the author get such details wrong. The sentences that imply such facts do not announce themselves explicitly as containing or implying 'facts to be believed'. Let's call sentences in fictions, which are of this kind, 'true-in-passing' (TIP) sentences.

⁵ Note that this is *not* the requirement that reader be able to infallibly identify *true* testimony from false. In virtue of its nature, testimony, and so beliefs based on it, may often be false.

If I'm right and, by including TIP sentences, authors don't ultimately intend to inform readers of truths about the world, then it might seem that we have opened the door to earlier worries. For if true TIP sentences aren't (I'm assuming) introduced deliberately and explicitly to inform readers, as in the case of testimony-in-fiction, then perhaps they're introduced for some other purpose which might equally well have been served by falsehood. And if this is right, two things seem to follow. One is that it might easily have been the case that the TIP sentences were false after all. The second is that it might easily have been case that the reader failed to identify likely-to-be-true TIP sentences and distinguish them from surrounding falsehoods. For now it seems that the usual signs of an authorial intention to inform can't be available to the reader to help her – and in the case of testimony-in-fiction, these were precisely what safeguarded against this latter sort of worry.

We can see these suspicions in action with respect to a sort of case discussed by Green. As he notes, fictions can represent, apparently accurately, facts about how things tend to be experienced, or appear (Green 2008). Authors can draw on their own real experiences to construct fiction, thereby apparently informing readers about those kinds of experience. For instance (my example), Irène Némirovsky drew upon her own experience of fleeing Paris to the countryside during World War II to inform her descriptions of Parisian evacuees in *Suite Française*. Through her prose we feel we learn something of the panic and desperation felt by those who are losing all that seems familiar; what it feels like to wait at a barricade, starving; or to be strafed by a plane. But we might still worry: if Némirovsky's intention, by the inclusion of these details, was not ultimately to inform, but to satisfy some goal equally well served by falsehood, then aren't the reader's subsequent beliefs formed lucky, even if true? Equally: might not the reader easily have missed the relevance of Némirovsky's description to the 'real world', given its 'in-passing' manner of presentation, and falsely believed something else from the fiction instead?

5 Understanding fictional genre as a safeguard against luck?

Green provides us with an ingenious-looking answer to such worries. Effectively, he offers an understanding of *fictional genres* as an additional means by which readers can reliably identify TIP sentences, and put epistemic faith in them, thereby eliminating the luck apparently involved.

According to Green, fictional genres are associated with norms about what subject areas or topics will be accurately described, in general terms, by works in those genres, and what will not. Authors can knowingly appeal to these norms to communicate true empirical claims to readers. For instance, he writes:

[I]t is part of the definition of a given genre to imply norms as to which departures from reality are permissible and which are not (2010, 356).

A detective novel requires rough accuracy about the ways of criminals and detectives. A political thriller needs to attain ballpark plausibility about governments and politicians, and so forth (*ibid.*).

A drama about a hospital emergency room needs to be reasonably accurate in its portrayal of the technologies and procedures available in that setting. By contrast, who comes into the E.R. with what gunshot wound or allergic reaction, and the family dynamics of these patients, are at the discretion of the screenwriter so long as she maintains fidelity to human psychology and the kinds of cases that present themselves in emergency rooms (2016, 295).

Given a particular genre and the filtering effect it has... readers may learn from literary fiction by accommodating material that passes through that filter (2016, 300).

If Green is right about this, then authors writing in the genres he mentions should normally be working hard to get the associated factual area in question right. Hence, perhaps, there won't be close possible worlds in which the author says something false about that area and the reader believes it. Equally, perhaps, we can dispense with the requirement that, for a knowledge-conveying sentence in fiction, there be a clearly discernible accompanying authorial intention to tell the truth; and replace this with a requirement that there be a clearly discernible accompanying authorial intention to respect certain genre norms. If so, presumably, readers can take this latter intention as an indirect sign that the author intends to tell the truth about a given factual area; where *this* gives her some reason to think that her sentences about this area are likely-to-be-true (see Section 2 above). Hence there won't be close possible worlds in which readers misidentify the likely-to-be-true sentences and believe something false instead.

Admittedly, all this is quite far from the terms in which Green presents his claims, but still, at first sight it promises to help enormously. On second look, however, the suggestion faces problems. For unfortunately, its crucial premise - that certain fictional genres have particular truth-telling norms associated with them - seems to be false; or at least, not true in the right way.

Take works in the genres referred to by Green: political thrillers, detective dramas, and medical dramas. Assuming we are talking about the most commonly-encountered populist versions of these, rather than avant-garde or otherwise specialist cases, it seems that in all three genres a central aim is plausibly construed as providing for the viewer or reader, an *immersive* experience. By this I mean, an experience in which she is deeply imaginatively engaged; one in which she can imaginatively 'lose herself' without much direct present awareness, perceptual or cognitive, of the non-fictional world around her. Immersive experiences are conducive to further presumed aims of the genres in question: to be *thrilling*, to be *gripping*, to be *diverting*, to *provide suspense*, or a *puzzle*, to be *arousing*, or to otherwise provide material for *fantasising pleasurably*. (The precise aim of a work will depend on the case in question).

One route to the goal of providing immersive imaginative experiences for viewers is the provision of a 'realistic' or 'plausible' fictional scenario⁶. But 'realistic' or 'plausible' fiction – and so fiction conducive to immersion – need not be the same as 'accurate' or 'true' fiction. What's important for imaginative immersion, and so for 'realism' or 'plausibility', is principally the reader or viewer's *not being made to notice disparities between the fictional scenario and the actual world* – for this would destroy her immersion. This means that inaccurate general descriptions anticipated by the author as identifiable as such by readers should be avoided. But on the other hand, if the factual area is one which, the author judges, the reader is likely to know little about in detail – as often the case with politics, crime detection, and medicine, in fact – then what matters for immersion is not accuracy, as such, but rather having 'convincing-sounding' detail included, which *for all the reader knows, might be true* (even if it isn't).

Now, one way to achieve 'convincing-sounding' detail, which, for all the reader knows, might be true, is to include actually true detail. As Ursula Le Guin puts it in her Introduction to *The Left Hand of Darkness* (of which more to come in Section 4 below):

[Fiction writers] may use all kinds of facts to support their tissue of lies. They may describe the Marshalsea Prison, which was a real place, or the battle of Borodino, which really was fought, or the process of cloning, which really takes place in laboratories, or the deterioration of a personality, which is described in real textbooks of psychology; and so on. This weight of verifiable place-event-phenomenon-behavior makes the reader forget that he is reading a pure invention, a history that never took place anywhere but in that unlocalisable region, the author's mind'. (Le Guin 2017 [1969], 11)

But crucially, including truth isn't the only way for a fiction writer to achieve such effects. If one is reasonably sure that readers won't know the difference, one can make detail up in a convincing-sounding way. And in fact, this is what authors of political dramas, detective novels, and medical dramas, among others, do a lot of the time. Hence the many internet articles by specialists working in the relevant fields, explaining to laymen what authors in fact got wrong about their fields. These articles miss the point: truth-telling for its own sake about the area in question was not the goal of the authors in question. Unlike in the testimony cases above, they are not to be reproached in the same way, if they get things wrong. However, on a charitable interpretation of authors working in immersive genres, what such articles do demonstrate is that those authors are not concerned with truth-telling for its own sake. They didn't *make mistakes*. They aren't stupid. They just didn't care about telling the truth about the detail in question, which is different.

⁶ Lack of 'plausibility' is a common complaint from readers, offered as a reason why they can't 'get into' a fiction.

The upshot is that works in any of the many genres whose primary aim is imaginative immersion, with a view to being gripping/ thrilling/ puzzling/ diverting/ arousing. etc. etc. do not in fact need accurate descriptive detail to be included to achieve that aim; but only descriptive detail which *appears to the uninformed to be true*. As acknowledged, one way of achieving this is to include accurate description, but it is not the only way. One might also include unidentifiably inaccurate description.

This apparently leaves the threat of luck hanging over many true beliefs the reader gets from reading TIP sentences in fictional works, even where those works fall into genres identified by Green. As yet we have no guarantee, as was perhaps initially promised, that authors working in the relevant genres must be working hard to get the factual area in question right. What they're probably working hard to do, rather, is to present detail which might for all the reader knows be true, and this isn't the same thing. Hence there still might be possible worlds in which the author says something false about that area, and the reader believes it instead. Nor are we in fact able to replace the requirement that there be a clearly discernible accompanying authorial intention to tell the truth via TIP sentences, with a requirement that there be a clearly discernible accompanying authorial intention to respect certain genre norms. For the latter does not in fact straightforwardly imply the former. Hence there may still be close possible worlds in which readers misidentify TIP sentences and believe something false instead. The worry about luck apparently remains.

Green might respond that he is concerned only with a *certain kind* of political thriller, detective or medical drama – one that explores some counterfactual, as described in Section 1. In these cases in particular, he seems to suggest, there's a special requirement on authors to get certain details right. For instance, he writes of an imagined political thriller, in which an innocent person is to be tortured to stop a terrorist from committing mass carnage:

What if a situation such as the former describes were to occur in a major city? That question, if it activates any genre at all, does not licence any departures from reality other than the following: The characters, political institutions, technological limitations, and the like need to be relevantly similar to those that we find in major, contemporary first world cities (2010: 358).

However, in actual fact, there seems no particular link between fictions that explore counterfactuals, on the one hand, and accuracy about detail within those fictions in a particular area, on the other. This is because the power of such fictions for readers, and why authors write them, tends to be to ask readers questions or imply conclusions about *general and widely applicable* hypotheticals. For instance, presumably the point of the thriller just described, is to ask whether readers would be willing to torture an innocent to save large numbers of lives. That general question can be raised effectively without particular concrete details in the work being accurate.

The best prospect for Green is a case where it is self-consciously and explicitly part of the definition of a genre that it will convey accurate information about some domain. For instance, as Green himself notes (2016, 285), the genre of 'hard science fiction' requires, as far as possible, accuracy about scientific laws and theories⁷. But this sort of case is the exception that proves the rule. Most genres aren't defined explicitly in terms of accuracy norms. Political thrillers, medical dramas, and detective novels certainly aren't.

6 An alternative proposal

Can we nonetheless show that true beliefs based on TIP sentences are in fact non-lucky? I think we can, for a limited range of them; and moreover, in a way that partially reinstates Green's claim that knowledge of genre can help identify what is true and false in a fiction.

Our earlier discussion contained the germ of a solution. That is, wherever the reader can non-accidentally ascertain that the authorial intention behind some descriptively accurate TIP sentence, in a particular case, refers to truth-telling in a way that cannot easily be eliminated, then *even though it does so only as an instrumental means to some further goal that isn't itself truth-telling*, we can still rule out either i) there being close possible worlds in which the author might have written something false, and the reader believed it; or ii) close possible worlds in which the reader misidentifies the true sentences and believes an accompanying falsehood instead.

Knowledge of genre can help here. For instance, as described earlier, some genres are aimed at imaginative immersion, with a view to being thrilling/ gripping/ arousing/ shocking/ etc. As noted, the precise goal will vary from genre to genre. In these genres, a central authorial intention behind a bit of fictional description is often, identifiably, to create a 'realistic' and so immersive world. In this case, a reader can at least assume that *any descriptive detail concerning entities that are easily checkable, or widely known about*, is in fact true. Where such details are accurate, there will be no close possible world in which they are false. Nor will there be close worlds in which the details are misidentified by readers, assuming readers understand the genre and its immersive aim, plus have a good working understanding of what, generally, readers will tend to know, and be thought to know by authors.

Equally, for some genres, there is an associated aim of creating nostalgia in the intended readership (for instance, works in the so-called 'Plantation tradition'; or certain 'nostalgia narratives' about the immigrant experience of exile). Here then, an author might non-accidentally include lots of true details about a given area, not to inform the originally intended readers of the work of new facts, but rather to remind them of facts which, the author assumes, they already know and recognise, with a view to evoking nostalgic feelings in them. Readers from a

⁷ The Historical Novel is another genre identified by Green as being defined in terms of accuracy norms (2016, 285). However there may be special problems identifying the true material in historical novels. See my 2017a.

much later period, or culturally distant readers, not themselves familiar with those facts, might well still understand the general aim implicit in the genre, and reasonably assume that any details which would presumably be known by the originally intended readers, are true; even if this truth is instrumental to the work's aims, rather than an ultimate aim in itself.

To sum up this proposal, then: the initial worry about luck with respect to true TIP sentences was based around a) the assumption that true TIP sentences are not included in a fiction primarily in order to tell the truth, but rather are included to satisfy some other goal; and the further assumption that b) TIP sentences which satisfy a) could easily have been false, and/or misrecognised by readers, assuming the goal in question was still satisfied. But, I have suggested, b) isn't always true, even if a) is true for TIP sentences. For there are some goals, in writing fiction, which require truth in particular areas, or at least make it by far the most sensible and least complicated choice⁸; and this fact is non-accidentally recognisable by readers.

This is a partial solution rather than a complete cure for the challenge from luck towards beliefs derived from TIP sentences. Some TIP sentences are likely to be included by authors in their fictions for reasons that have nothing to do with truth, even instrumentally; or which contain an eliminable reference to it. For instance, as already noted, if the aim is immersion, and the factual area described is one the author judges that her intended readership won't know very much about, she is free to make things up, even if as a matter of fact she includes true detail because she can't be bothered inventing. True beliefs acquired by the reader in this case would be lucky. Or: if an author witnesses an unlikely and hilarious event in real life, and she then accurately describes it in a fiction simply in order to provide an unlikely and hilarious plot point, the description's truth will be completely incidental to its inclusion, and so any true beliefs the reader got as a result of reading about the event would be lucky. But actually, this looks like the right result to me: such beliefs *would* be lucky! So what we seem to have arrived at is a useful way of distinguishing knowledge derived from TIP sentences in a fiction, from merely lucky true beliefs derived from TIP sentences, which don't yet count as knowledge.

7 Counterfactual fictions and the challenge from luck once again

I turn now to the category Green is most centrally concerned with in his writings: that of fiction that supposedly produces knowledge of counterfactual truth. I'll first focus here on 'didactic' fictions, as Green calls them. Is there a threat of luck here too? I'm afraid I think there might be.

Recall that with 'interrogative' fictions, in Green's terms, effectively the reader can hold the whole of the content of the fiction fixed, as a supposition, and then

⁸ Stacie Friend writes of a reader possibly 'developing a strategy that works very well' for a particular genre, which allows her to 'by and large ... believe only what is true' (2014, 244).

ask herself – if *that* was all true, what would happen?/ how would I feel?/what would I do? This means, obviously, that these fictions, as such, are not making claims in any sense, but merely asking questions. If the reader arrives at (self-) knowledge as a result of reading such fictions, she has not done so by accepting any claim implicit in the fiction, but rather via a process of reflection on her own. In contrast, in the case of a ‘didactic fiction’ in Green’s sense, the reader is being asked by the author to hold a certain amount of the fictional content fixed, as ‘what is supposed’ (i.e. in philosophical terms, to treat it as ‘antecedent’) and to treat the rest of the content as implying a claim about what would follow from the original supposition (i.e. to treat it as ‘consequent’). So we are to treat the fiction as a whole as a counterfactual claim: if *that* happened, then *this* would/could happen.

I actually doubt that very many works are unambiguously ‘didactic’ in this sense. Take an example Green himself discusses: Ursula Le Guin’s *The Left Hand of Darkness*. Green presents this novel as one which concretely represents the consequences, were people able to change gender throughout their lives (2010, 355; 2017, 54). But in fact Le Guin seems to have been quite ambivalent about whether, in this book, she was only asking a question, or also giving an answer; and if the latter, what answer exactly she was giving. Admittedly, in places, she says things that look rather helpful to Green’s cause. For instance, in her original Introduction she writes:

If you like you can read [the book].., and a lot of other science fiction, as a thought-experiment. Let's say (says Mary Shelley) that a young doctor creates a human being in his laboratory; let's say (says Philip K. Dick) that the Allies lost the second world war; let's say this or that is such and so, and see what happens.... (2017 [1969], 10).

And in a later essay ‘Is Gender Necessary?’ she writes:

[In *The Left Hand*..] I eliminated gender to find out what was left (1989, 160).

However, in other places, even within the same essay, she is much more tentative.

[A]s an experiment, it was messy. All results were uncertain; a repetition of the experiment by someone else, or by myself seven years later, would probably give quite different results’ (1989, 160).

[Thought experiments] .. are questions not answers; process, not stasis. One of the essential functions of science fiction, I think, is precisely this kind of question-asking. {1989, 159; see also 171-2}.

Generally, it seems to me that Le Guin’s ambivalence about the precise aim of this book – is to formulate a specific detailed counterfactual, or merely a very general and vague one? Or indeed, is it only to pose questions about alternative

states of affairs but not answer them?– undermine any confidence one might have that it is clearly didactic in Green’s sense.

Or take another example of Green’s, Aldous Huxley’s *Brave New World*. As already noted, Green takes Huxley effectively to be saying in this work that ‘were society to exist based on principles associated with hedonic utilitarianism’ then ‘the State would support an orgiastic religion, drug-induced anaesthetization of its citizens, suppress all independent inquiry, and so forth.’ (Green 2010, 360.) Green then adds that ‘The reader is then justifiably expected to recoil with horror at such a world’ (*ibid.*). Yet in his 1946 Foreword to a new edition of the work, Huxley says that ‘in *Brave New World* this standardization of the human produce has been pushed to fantastic, though not perhaps impossible, extremes’ (1994, 31). This suggests to me that he could not have been claiming anything as determinate as Green suggests. Meanwhile, according to Huxley scholar David Bradshaw:

[I]t seems more than likely that the composition of *Brave New World* proved so problematic for Huxley between April and August 1931 because he was unsure in his own mind whether he was writing a satire, a prophecy or a blueprint (Bradshaw 1994, 7).

Here too, it is not clear that Huxley was certain about what he firmly wanted to conclude from the circumstances he was fictionally exploring.

So, based on this partial sample at least, perhaps the class of unambiguously didactic fictions is smaller than it first looks. In any case, assuming that there are at least some, we can still ask: for a ‘didactic’ fiction which implies a true counterfactual (and let us also assume that some do): are there close possible worlds in which the same counterfactual is implied by the author but is false and the reader still believes it? Or: are there close possible worlds in which the reader misidentifies some other counterfactual as likely-to-be-true instead? I fear there are both.

One worry here pertains to author expertise. For the claim that fictions can give us knowledge about counterfactuals to be interesting, it has to be more than knowledge of bare possibilities, in the sense of conjunctions of events that are *not impossible* (see also Green 2010, 352). For cognitive value, we must be talking about counterfactuals that are somewhat likely or probable; which ‘could well happen’. Now, whereas an author can easily study a particular area of empirical fact, it’s harder to see what they can do to make themselves an authority about probable counterfactual truths in a given area. The question especially arises for the relatively complicated counterfactuals that – we are assuming – are represented in some fictions. It is often easy enough to work out that if *x*, then *y*; it’s much harder to work out that if *a, b, c, d, e, f* and *g*, then *h, i, j, k* and *l* would follow. It’s unclear what background would differentiate an author’s reliably getting them right rather than accidentally hitting the mark.

Green might reply that we don’t particularly need evidence of authorial expertise about counterfactuals, for the truth of a counterfactual is substantiated by its

seeming 'compelling' to readers, and progressing 'in a way that seems inexorable or at least highly plausible' (2017, 57). But this faces a familiar worry: if the area is one that the reader herself knows nothing much about, then what seems compelling needn't be the same as what is true.

A different worry pertains, once again, to what the reader, rather than the author, might easily have done differently. This worry is that, in the already narrow class of 'didactic' fictions, it is very often unclear to the reader, for general structural reasons, what counterfactual the author intends to endorse, exactly. To go back to the counterfactual supposedly present in *Brave New World*: what would make *this* the identifiably likely-to-be-true counterfactual, rather than, say, *were society to exist based on principles associated with hedonic utilitarianism, AND the State was to support an orgiastic religion AND drug-induced anaesthetization of its citizens, THEN the State would suppress all independent inquiry*? Or, in general terms: how is the reader to non-accidentally distinguish what's intended to be the initial supposition (the antecedent), and what's intended to follow from that initial supposition (the consequent)?

There's an important contrast to be made here with thought experiments explicitly used for cognitive purposes, in academic domains such as philosophy, physics, or economics. These typically differ in at least two ways from novels and stories. In the first respect, a thought experiment in an academic context is usually, at least on the face of it, interrogative in form: it doesn't spell out (at least, as part of the content of the experiment itself) what intended conclusion is supposed to be drawn by the reader. This means that we can read it according to an implicit convention of interpretation: whatever is explicitly described in the text is to count as antecedent, to be supposed; whatever the reader works out as following from what is explicitly described, counts as consequent (assuming, at least, that the reader's reflections correspond to what the author intends her to think). Secondly, the surrounding context in which an academic thought experiment is presented – usually, a text, arguing explicitly for a particular conclusion – will nonetheless unambiguously identify the conclusion the reader is expected to draw from the experiment, as she reads it. Should this explication, unusually, be lacking for some reason, there will still remain the general context of the text itself, and its identifiable purpose, to help orient the reader. Appreciating this context and purpose will help the reader to work out why the thought experiment was introduced, and what it is supposed to show.

Most 'didactic' novels and stories aren't like this, I suggest. Admittedly, it is easy enough to identify the central 'premise' of a 'didactic' fiction, because of its obvious interest, at least to the intended readership. But still, even so, lots of the accompanying fictional detail will then surely be ambiguous between a) what is supposed to follow from that initial supposition b) what is being added in as *new* supposition, and not particularly intended to follow from anything already stated. Any accompanying claims of the author will be highly unlikely to spell out the difference between these two groups, exactly, I would think. I take it to follow that true beliefs about counterfactual states of affairs, derived from 'didactic' fictions, and which go beyond platitudes into some detail rather than simply being vague, are very often going to be lucky in our extended sense. The

truths in question will only be accidentally identified by readers as such, and so accidentally believed.

8 Interrogative fictions: a better bet

True beliefs acquired by readers from 'didactic' fictions look problematic, and often won't count as knowledge, I have suggested. When we turn to 'interrogative' fiction, however, things look a lot more cheerful. For none of the points just made apply here.

Many fictions are clearly intended to be taken by readers as 'interrogative', and are understood as such. The novelist Jodi Picoult, for instance, has made it her calling card to describe morally challenging scenarios for fictional characters, and so effectively ask difficult moral questions of the reader along the lines of 'what would you do if...?'. This effect is enhanced by editions marketed specially for home book clubs, which often end with a list of questions asking how the reader would respond, were she in the situation of the character in question.

In works of this kind, moreover, crucially, there is an easily detectable split between antecedent and consequent; for, as in a traditional thought experiment, what is explicitly stated in the text is clearly to be understood as stipulated and antecedent; and what the reader then concludes, upon reflection, is to be understood as consequent, assuming of course that she reasons well. Equally, there is no particular worry about authorial expertise this time: for the author isn't stating anything, but only asking a question.

A question remains, I suppose, about any true counterfactual conclusions drawn by readers in response to interrogative fictions: are they accidentally true? Here too former worries about expertise might bite, especially if the counterfactual concerns some complicated empirical projection, or difficult moral question. However, with a particular kind of conclusion, at least, our ground is surer. This is what Green calls a conclusion of 'de se supposition' (2010, 362): working out what I, the reader, would do or feel or think, were some fictionally depicted state of affairs the case. On the assumption that my imagining certain circumstances gives me some good (though defeasible) evidence about what I would then do, feel, or think in those circumstances, then any resulting true beliefs in this area look grounded in evidence, and recognisably so, to an extent which renders them non-lucky⁹. This assumption is a central pillar of Simulation Theory in the Philosophy of Mind, and I see no reason to doubt it.

It remains only to sum up what I've argued in this paper. In order for true beliefs acquired from reading fiction to count as knowledge proper, they must survive what I've called 'the challenge from luck'. That is, it must be established that such beliefs are neither luckily true, nor luckily believed by readers. I considered three kinds of true belief a reader might get from reading fiction: a) those based

⁹ My claim is not the strong one that one always gets true beliefs from such a process; but rather that, where one does get true beliefs, they are not luckily true.

on testimony; b) those based on what I call 'true in passing' sentences; and c) those beliefs about counterfactuals one may get from reading a 'didactic' fiction. The first group escaped the challenge from luck relatively easily, but things turned out to be more complicated with the second group. For that group, I examined the suggestion that knowledge of *fictional genre* may see off the challenge from luck. However, I rejected this suggestion, adapting it substantially instead to offer beliefs of this kind a more promising escape route. Finally I discussed the category of 'didactic' fiction. I argued that any true beliefs about counterfactuals gained from such fictions are likely to be lucky; but that things are much more promising for any 'de se' true counterfactual beliefs gained as a result of engaging with 'interrogative' fiction.

Many thanks to Tony Booth, Manuel Garcia-Carpintero, and an anonymous referee for helpful comments.

Bibliography

Adler, Jonathan (2017). "Epistemological Problems of Testimony". In *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2017/entries/testimony-episprob/>>.

Bradshaw, David (1994). "Introduction to Aldous Huxley's *Brave New World*". Flamingo Modern Classics.

Davies, David (2007) *Aesthetics and Literature*. London: Continuum.

Deutsch, Harry (2000) "Making up stories". In *Empty Names, Fiction and the Puzzles of Non-Existence*. Eds. A. Everett and T. Hofweber. Chicago: University of Chicago Press pp.149-81.

Fowler, Karen Joy *We Are All Completely Beside Ourselves*. London: Serpent's Tail.

Friend, Stacie (2014) "Believing in Stories". In Greg Currie, Matthew Kieran, Aaron Meskin, and Jon Robson (eds.) *Aesthetics and the Sciences of Mind*. Oxford: Oxford University Press, pp. 227-48

Green, Mitchell (2008) "Empathy, expression, and what artworks have to teach". In G. Hagberg (ed). *Art and Ethical Criticism*. Oxford: Wiley Blackwell. pp.95-122.

Green, Mitchell (2010) "How and What Can We Learn From Fiction?". In G. Hagberg and W. Jost (eds.) *The Blackwell Companion to Philosophy of Literature*, pp. 350-66. Oxford: Wiley Blackwell.

Green, Mitchell (2016) "Learning to be Good (or Bad) In (Through) Literature". In G. Hagberg (ed.) *Fictional Characters, Real Problems: the Search for Ethical Content in Literature*. Oxford: Oxford University Press. pp. 282-305.

Green, Mitchell (2017) "Narrative fiction as a source of knowledge". In P. Olmos (ed.) *Narration as Argument*. pp. 47-63. New York: Springer.

Huxley, Aldous (1994 [1946]) "Foreword". *Brave New World*. Flamingo Modern Classics.

Le Guin, Ursula (2017 [1969]) "Introduction". *The Left Hand of Darkness*. Gollancz.

Le Guin, Ursula (1989) "Is Gender Necessary? Redux". In her *The Language of the Night, 2nd Edition*. Women's Press.

Pritchard, D. "Anti-Luck Epistemology". *Synthese* (2007) 158: 277.

Stock, Kathleen (2017a) "Fiction, testimony, belief and history". In: Sullivan-Bissett, Ema, Bradley, Helen and Noordhof, Paul(eds.) *Art and Belief*. Mind Association Occasional Series . Oxford: Oxford University Press. pp.19-42.

Stock, Kathleen (2017b) *Only Imagine: Fiction, Interpretation and Imagination*. Oxford: Oxford University Press.