

Absorbing Reality into Fiction: The Challenge of Reading Fiction with Reality in Mind

Ben Martin (University of Bergen)[†]

The truths of a fiction extend far beyond those stated in its text. We have good reason, for example, to believe that Sherlock Holmes possesses a Cerebral Cortex, although we are never told so within the stories. To infer these further truths, we must rely upon a host of previously acquired background assumptions. Establishing what exactly these assumptions are, and how they help us extend the truths of a fiction, however, is not a simple matter. Contributing to the debate, Stacie Friend (2017) has recently argued that fictions fundamentally rely upon the actual world for their content, appealing to evidence from the cognitive sciences. According to her Reality Assumption, every proposition which is actually true is also fictionally true, unless otherwise excluded by the fiction. This paper challenges Friend's Reality Assumption, arguing both that the empirical evidence advanced in its favour offers inadequate support, and that the Assumption in its present form inadequately serves its intended function.

Keywords: Fictions, Truth-in-fiction, Reality Assumption, Reality Principle, Friend.

1. Identifying Truths in Fiction

Doyle never tells us that Holmes lives nearer to Oxford than Edinburgh, yet if asked we would be certain he does. Similarly, Doyle never rules out the possibility that Holmes has a third eye, yet if pushed we should flat out deny he does. Given that no sentences in the Holmes stories themselves warrant these claims, nor are they logically entailed solely by what is said textually, we may rightly wonder what licenses such purported knowledge about the unseen world of Sherlock Holmes.

We could begin by drawing an analogy with everyday conversations. We take shared background information into conversations, drawing inferences from a combination of these assumptions and new information our fellow conversers introduce, and reading implications into the content using conversational norms. If I find myself in an unfamiliar town and ask a local where the nearest restaurant is, it isn't necessary for her to tell me the exact street. Her assertion that 'The main drag is about a five-minute walk from here', pointing to the east, suffices. By our shared understanding of the conversational norm that any contribution should be *relevant*, and our shared background knowledge that a town's main drag often contains

[†] Contact: Benjamin.martin@uib.no.

restaurants, the local successfully communicates with her assertion and gesture that the nearest restaurant is a five-minute walk to the east. Initiated by Grice's (1975) pioneering work, these conversational processes are now reasonably understood. Why, then, can we not simply transmit these conversational principles to our understanding of fictions? After all, we might well think of works of fiction as very one-sided dialogues between an author and her readership. I can deduce from what Doyle tells me that Holmes lives nearer to Oxford than Edinburgh, knowing both that Baker Street is in London and Oxford is closer to London than Edinburgh. Similarly, Doyle is not required to say that Holmes doesn't have three eyes; it's a shared standing assumption that humans have at most two.

In fiction, however, the matter just isn't so simple. For multiple reasons the principles of conversational implicature do not easily transfer to the reading of fiction. Firstly, not all fictions are written by our contemporaries. Consequently, we shouldn't expect to share all of our background assumptions with the author. Reading non-contemporaneous fiction can be like having a conversation with one from a foreign and unfamiliar land. Nor does it take much to shake our confidence that the author shares our background assumptions. Homer's description of the sea as "wine-dark" is enough to make us doubt how many assumptions the author and we share, blocking inferences and implicatures. After all, if the author willingly admits that the sea is red in colour, when it is obviously a greenish-blue, what other strange and wondrous assumptions underpin the fictional world?

Secondly, the purpose of many fictions is to remove commonplace background assumptions, introducing enlightening counterfactuals. In this context, it is not suitable to simply plug all of our background assumptions into our reading of the story. Nor can we wait for the author to explicitly remove them, replacing each with new assumptions suitable to the new fictional world, for there are simply too many. By means of replacing them, then, we must do our best in a piecemeal fashion, hypothesizing which of our normal assumptions are most likely to be discarded, and adapting these levels of credence as we progress, like any good Bayesian. At times, we may even need to entertain multiple worlds as *the fictional world* at any one time, just as a scientist may informally entertain multiple possible explanations for a phenomenon, allowing for the possibility of further evidence to differentiate between the candidates.¹

¹ While it's not unusual to speak of the content of fictions in terms of *fictional worlds*, we should not equate these fictional worlds with the worlds of possible-world semantics. The latter worlds, at least in *normal* possible-world semantics, are both consistent and complete, yet fictional worlds are (almost) never complete and sometimes inconsistent. Doyle never tells us, for example, the colour of Watson's eyes, but does tell us that Watson has one bullet-wound, which happens to co-exist in his shoulder and leg.

Lastly, unlike everyday conversations, where we assume that the other party isn't attempting to fool us, the use of unreliable narrators calls into question whether we can simply take what is explicitly stated in a text at face-value. It is possible that the author has created the narrator as either fallible, and thus flat out wrong about certain of the claims made, or worse a liar, purposively leading us astray within the fiction. Thus, complications arise not only when drawing inferences to further fictional truths from the text, but in identifying which portions of the text themselves constitute fictional truths.

All three of these factors make providing a successful completion of the *truth-in-a-fiction* schema,

Some proposition p is true in a fiction f iff p possesses properties P_1, P_2, \dots, P_n

particularly challenging. To do so, we must establish both which propositions are *directly* generated by the fiction, and which are *indirectly* generated. That is, which truths we can draw directly from the source, such as the text, known as the fiction's *primary truths*, and which we can only appreciate by making inferences from other fictional truths, the fiction's *secondary* (or, *implied*) *truths*.

It should be no surprise that there is a lack of consensus on how to identify either type of truths within a fiction. While it's clear we cannot assimilate *primary* truths with what is explicitly stated in a text, both because of authorial mistakes and unreliable narrators, what is needed to supplement such textual evidence is not. Genre considerations would certainly seem to play a role. We have very good *direct evidence* from a fiction being a vampire story that some of its characters drink blood, even if the text resists mentioning this fact. It is such genre tropes that allow for authorial subtlety and playfulness in a genre over time – just think of the evolution of the sonnet. What else, though, needs inclusion within these *primary* truths? It seems we often require the assumption of certain linguistic and cultural conventions from the author's community to understand complex matters of reference, metaphor, and audience expectations. For example, that calling someone a 'cat' in a jazz club doesn't mean they are feline, and that a Mafioso delivering a kiss to another need not be a sign of affection. However, we then face the seemingly intractable problem of deciding which of a society's conventions we are bound to include in the fiction's *primary* truths, and which we can avoid. These are complex matters, and the literature so far has merely scraped the surface of the subtlety with which readers must draw information out of a text.

In contrast, the philosophical literature is full of attempts to account for how *secondary* truths are generated, and the complexities that arise from such accounts. By completing the *generation schema*,

If propositions p_1, p_2, \dots, p_n are true in some fiction f , then some other proposition, q , is also true in f iff q has the relation(s) R_1, R_2, \dots, R_n to p_1, p_2, \dots, p_n ,²

and explaining the exact relationship q must have to p_1, p_2, \dots, p_n to be true in the fiction, these accounts hope to specify *all and only* those truths which it is reasonable to include within a fiction. These *principles of generation* range from whether it's reasonable for an informed reader of the text to infer that the fictional author of f believed q on the basis of p_1, p_2, \dots, p_n (Currie 1990), to whether the (ideal) reader could infer that the (idealized) author is inviting the reader to make-believe that q on the basis of p_1, p_2, \dots, p_n (Byrne 1993), and whether members of the author's community would mutually believe that q on the basis of p_1, p_2, \dots, p_n , otherwise known as the *Mutual Belief Principle* (Lewis 1975; Walton 1990: Ch. 4).

While each of these principles has its strengths, each has substantial weaknesses. None of the principles, for example, seem to be able to adequately account for why we are justified in believing that Achilles has a Cerebral Cortex in the *Iliad*, or that being discriminatory towards someone, or their culture, on the basis of their skin colour is morally wrong even in the *Heart of Darkness*. It is the perceived inability of these principles to explain how we are able to integrate our understanding of the real world into our appreciation of a fiction that accounts for the continued intuitiveness of the most famous of the *principles of generation*, the *Reality Principle* (Lewis 1975):

If propositions p_1, p_2, \dots, p_n are true in some fiction f , then some other proposition, q , is also true in f iff were it actually the case that p_1, p_2, \dots, p_n , it would be the case that q .

The *Reality Principle* asks us the reader to use the actual world as our guide to whether q ought to be included as a truth of f or not. We must assume that the actual world includes all of the truths already recognised within the fiction, and ask whether in such a scenario q would also be true. By considering the relevant counterfactuals, and answering in the positive, we admit q as true in the fiction.

² Note, p_1, p_2, \dots, p_n need not be *primary* truths of f . Secondary truths can just as easily be generated from other secondary truths.

Through integrating our understanding of the actual world into fictions, the *Reality Principle* is able to account for the existence of Cerebral Cortices in the *Iliad* and the immorality of racism within the *Heart of Darkness*. Yet, the principle still has weaknesses. These include, but are in no way restricted to, its inability to account for *impossible* stories, whose propositions *could not* be integrated into the actual world, and its inability to account for reasonable interpretations of a fiction based upon what are now quite esoteric beliefs on the part of the author and their community, such as in gods and ghosts. The *Reality Principle* will, uncharitably, always mandate the most *realistic* reading of a text to modern ears and not necessarily the most aesthetically rewarding.³

The fact that each *generation principle* suffers significant shortcomings may reasonably lead one to being sceptical about the possibility of detailing the exact conditions under which a proposition should be deemed true within a fiction. Walton (1990: 139) admits as much by suggesting there is no single, successful, unifying *generation principle*. Yet, even if we agree with Walton, and conclude that no one *generation principle* can provide the correct and complete implied truths of *every* fiction, we may still wonder what informative claims *can* be made about how we establish the *secondary* truths of fiction. After all, one does not need to establish a set of necessary and sufficient conditions for a phenomenon to say something informative about it.⁴ We might hope to discover certain background assumptions that any reasonable *generation principle*, regardless of their idiosyncratic consequences in certain situations, must recognise. It is just such a project that Stacie Friend (2017) has recently taken up in arguing that, regardless of which *generation principle(s)* are ultimately to be used in a given circumstance, all should presume that “everything that is (really) true is also fictionally the case, unless excluded by the work,” (Friend 2017: 29). Unsurprisingly, she calls this presumption the *Reality Assumption* (RA).

2. The Reality Assumption

In advocating the RA, Friend is clear she is not attempting to provide a complete account of how *secondary* truths are generated within a fiction. Rather, her goal is far more modest, to provide just one piece of the puzzle. The RA “does not tell us how to make such inferences [to the secondary truths]”, but rather “supplies an array of story-truths...which, in combination

³ For an overview of the strengths and weaknesses of the *Reality Principle*, see Walton (1990: Ch. 4) and Woodward (2011).

⁴ An excellent example is Williamson’s (2000) analysis of knowledge.

with other storified content, can form the basis of inferences to further story-truths,” (Friend 2017: 32). Thus, we should see Friend’s RA as providing a substantial number of inputs for our *generation principles*. If these principles are our rules of inference when navigating fictions, then the RA supplies our (defeasible) starting points.

However, in another sense Friend’s proposal is not so modest, for she does not think of the RA as supplying one of many possible starting points for our reading of fiction. Rather, it “specifies a comprehensive default position *for which there is no alternative*,” (Friend 2017: 34. My emphasis). Consequently, when evaluating Friend’s proposal, we must remember the strength of her claim – there is *no other plausible starting assumption* to take when it comes to reading *any fiction*.

To fully appreciate the RA, two clarifications are necessary. Firstly, the RA requires not that *our own beliefs* about what is true in the actual world should be our starting assumptions when deciding what is true within a fiction, but *what is actually true*:

The Reality Assumption has several surprising consequences, such as that fictional truths go far beyond what authors or readers could know. (Friend 2017: 29).

[The RA instructs us] to take for granted that *everything that obtains in reality* is storified, adjusting this presumption only as needed. (Friend 2017: 34. My emphasis)

According to the Reality Assumption, *facts* about the real world are storified...by default. (Friend 2017: 34. My emphasis).

It is hard to make sense of these claims unless it is actual truths which are the RA’s content, and not our *beliefs* about the actual world.⁵ Beliefs are not facts, nor do our beliefs always correspond with reality. Additionally, it only makes sense to say the RA will ensure that fictional truths go beyond what we could know if these assumptions are not constituted wholly by our own beliefs. Thus, it is the truths of the actual world that RA requires we defeasibly take for granted when reading fictions, and not our beliefs about the world.

Secondly, in accordance with Walton’s (1990: 40) *Pretence theory*, Friend (2017: 30) understands fictionality in terms of “what the work invites imagining,” with an *invitation* to imagine understood as a conditional prescription:

⁵ Additionally, as we shall see below, we can only make sense of Friend’s criticisms of competing assumptions if we interpret the RA in this fashion.

We are *invited* to imagine p iff, if the question arises and we must choose between imagining p and imagining not- p , we are required to imagine p .

Thus, what we are *invited* to imagine when engaging with a work need not be what we actually imagine, either because the question of which proposition we ought to imagine hasn't yet arisen, or because we have simply disobeyed the work's requirement to imagine p . Additionally, what a work *invites* us to imagine doesn't exhaust what is *permissible* for us to imagine. For example, a work will often fail to comment on a character's eye colour or handedness. In such cases, we are not *invited* to imagine the characters in any particular way in this regard, yet we are certainly *permitted* to fill in the details.

We are now in a much better position to provide a fuller articulation of the RA:

For any fiction f , every proposition p that is true in the actual world, we are invited to imagine as part of f (i.e. p is *true in f*), unless p is excluded (in some sense) by another *primary* or *secondary* truth in f .

According to Friend, unlike past failed attempts to provide a reasonable *generation principle*, the RA doesn't fall foul of any charges of incompleteness or implausible consequences. After all, the assumption only provides a necessary condition for fictionality, and thus charges of incompleteness would be wholly inappropriate. At the same time, the RA possesses the benefits of the *Reality Principle*, explaining why we are right to insist that Achilles has a Cerebral Cortex, and why racism is morally wrong in the *Heart of Darkness*, regardless of the views of narrators or their contemporaneous audiences. Thus, in contrast with the *Reality Principle*, "the Reality Assumption applies universally" (Friend 2017: 34).

In what follows we take issue with Friend's proposal, suggesting that, firstly, evidence for choosing the RA over competing assumptions is relatively weak and, secondly, the proposal suffers from certain substantial challenges.

3. Arguments for the Reality Assumption

Friend's positive case for the RA relies upon findings from cognitive psychology. According to Friend (2017: 34-5), these findings provide strong evidence for the claims that, firstly, "our comprehension of stories relies on the Reality Assumption," and, secondly, that this reliance is automatic and unreflective in "the ordinary course of reading."

In support of the first claim, Friend highlights research suggesting that our experiences in the real world prime certain readings of fictional situations, and findings which demonstrate participants are often willing to transfer real-world truths into fictional scenarios. In a study by Gernsbacher, Goldsmith and Robertson (1992), participants were given several short stories to read, with each presenting a protagonist involved with a secondary character. One of these stories relayed a character Tom stealing from his best friend Joe's workplace, with rough consequences for Joe:

Joe worked at the local 7-11, to get spending money while in school. One night, his best friend, Tom, came in to buy a soda. Joe needed to go back to the storage room for a second. While he was away, Tom noticed the cash register was open. From the open drawer Tom quickly took a ten-dollar bill. Later that week, Tom learned that Joe had been fired from the 7-11 because his cash had been low one night.

At the end of each story, the experimenters included a target sentence containing an emotive word. Two versions of the target sentence for each story were produced, one containing an emotive word considered suitable to the context, and the other an unsuitable emotive word. For example, with the story above the following two target sentences were provided:

- (1) It would be weeks before Tom's *guilt* would subside.
- (2) It would be weeks before Tom's *pride* would subside.

Across all the stories provided, target sentences were read considerably more slowly when containing emotive words mismatched to the scenario, implying that participants' real-world knowledge of folk psychology primed their reading of the target sentences, informing their understanding of the stories.

Further support for the claim that participants import real-world information into their imagining of fictional situations comes from the Weisberg and Goodstein (2009) study. Participants were given three stories with the same basic narrative: a boy and his dog went to a zoo and rescued another child's kite from flying away. Each story, however, differed in the number of actual-world facts they violated. In the 'Close story' no obvious violations of reality were mentioned, whereas in the 'Middle story' the boy was said to have the characteristics of a superhero (e.g., wearing a cape, having resistance to pain, and the ability to fly), and in the 'Far story' the boy additionally had the ability to teleport and shapeshift into non-human animals, while his dog could speak to him and the zoo contained aliens. Participants were then given a list of twenty-one real-world truths, nine of these were contradicted within the Middle

and/or Far stories, and twelve others had no connection to the content of any of the three stories. These twelve unconnected real-world truths fell into four categories:

Contingent (E.g., “Washington, DC, is the capital of the United States”)

Conventional (E.g., “It is considered rude to pick one’s nose”)

Scientific (E.g., “People have hearts”)

Mathematical (E.g., “Two plus two equals four”)

When asked whether they believed each of the real-world truths were “definitely true”, “probably true”, “probably not true” or “definitely not true” in each story, participants rated those real-world truths which had been contradicted as much less likely to be true in both the Middle and Far stories than those not contradicted. In addition, it was found that participants were more likely to admit the truth of the non-contradicted real-world facts in the Close story than in the Middle and Far stories. Further, participants were more likely to admit the continued truth of the Mathematical real-world truths in all three stories than the Scientific real-world truths, which were themselves more likely to be admitted as true in the stories than Conventional truths, with Contingent truths being the least likely to be deemed true in all three stories. These results suggest that the real-world truths included within the fictions are impacted by at least two factors: i) The *distance* of the fiction world from the real world; ii) How *integral* the real-world truths are to our representation of the actual world. Thus, while simple arithmetical truths were more likely to be included in a story than social conventions, the likelihood of a social convention being included depended (at least) partly upon how different in other respects the story was from the actual world.

Friend (2017: 36) takes both of these findings to provide significant evidence for the truth of the RA:

The empirical evidence shows that our ability to make inferences that take the real world as background is essential to our capacity to understand what we read. Even when we exclude certain real-world facts, we still take for granted that others are storified.

Additionally, findings from several studies suggest that priming effects, such as those observed in Gernsbacher *et al.* (1992), are both stable across differing levels of comprehension and spatial imagery abilities (Haenggi, Gernsbacher and Bolliger 1994), and still present even when participants are required to perform a divided-attention task (Gernsbacher, Hallada and

Robertson 1998). Both findings constitute evidence that these priming effects are automatic ‘off-line’ responses, rather than the product of reflection.

How significant a support, then, do these findings lend the RA? Firstly, it’s important to remind ourselves what the RA requires of us. It dictates that unless somehow excluded by *f*’s other truths, we are invited to imagine as part of *f* every proposition which is actually true. Nor is this just one of many possible starting points for reading fictions, it is the *only* plausible starting point. Thus, for these studies to support the RA, the *best possible explanation* for their results must be that fictions invite us to import real-world truths into our relevant imaginings.

Before we consider whether the RA is the most plausible explanation of the results, let us first recognise three complications that arise from using the studies’ results as evidence for the RA. Firstly, Friend’s RA is intended to be the *only* suitable possible starting point for engaging with a fiction, that “[taking] the real world as background is essential to our capacity to understand what we read,” (Friend 2017: 36). Yet, it’s unclear how any of the empirical evidence Friend appeals to could demonstrate this. Certainly, this evidence could show that we do, and are *best off*, in the cases considered within the studies engaging with the passages from a background of the real world. However, the stories considered within the studies are vignettes at best, shorn of authorial richness, subtlety and ambiguity. It would then be hasty to extrapolate from these cases to the claim that the RA is *only suitable* backdrop to reading *any story*. In order to draw such a conclusion, we ought to test this hypothesis directly against other possible candidates for such a backdrop, such as the *Mutual Belief Assumption*, which we will discuss later, using a more diverse sample of works admitting of various interpretations. Only then can we say with any confidence that the empirical evidence shows that the RA is the *essential* background when engaging with fictions.

Secondly, Friend takes this empirical evidence to show that we *should* use the real world as our background when engaging with any fiction, and not only that we tend to do. However, for the same reasons as above, the empirical evidence Friend proposes in favour of the RA cannot demonstrate that we *ought to* engage with fictions in this manner. To show this, we need reasons to believe that no other initial assumption *could* yield such fruitful readings of a text. Friend is aware of this criticism herself:

Critics may retort that [the empirical results support] merely a descriptive claim. The fact that we *do* rely on the Reality Assumption does not entail that we *should*. Perhaps readers are systematically mistaken in deploying the Reality Assumption. But such an error theory is implausible. The basic skill of understanding stories is not exclusive to the erudite; it is common among children, as well as across cultures.

But, the prevalence of a norm does not demonstrate its fruitfulness. Many individuals' actions are a result of *loss aversion*, for example, although these actions are often damaging. Thus, it's completely possible to show that a common practice should not be endorsed by introducing reasons for why it's suboptimal. As we shall argue in the next section, this is exactly the position the RA finds itself in.

Lastly, at present there is a lack of clarity within Friend's articulations of the RA over what it means for a fiction to 'exclude' real-world truths. Does this mean something as strong as logical or metaphysical exclusion, or as weak as other fictional truths making real-world truths unlikely? Without a clear answer, it's impossible to make precise predictions on the part of the RA to compare with results from empirical studies. We cannot be sure, exactly, when the RA ought to insist that a reader import a certain real-world truth into her representation of some fiction, and when not. As shall see, ambiguity over the meaning of 'exclusion' causes not only complications when searching for positive evidence for the RA, but undermines its plausibility.

It is now time to consider whether the available empirical data adequately supports the RA, rather than a competing assumption. As it stands, the RA dictates that what ought to be imported into our imaginings of a fiction are not our *beliefs about the actual world*, but those propositions which are *actually true*. Now, presumably, our set of beliefs about the world isn't identical to the set of actual truths. We are fallible and our beliefs are incomplete. Thus, in order for the empirical evidence to support the RA, we must have reason to believe the evidence fits better the RA over a competing hypothesis, which we shall call the *Reader Belief Assumption* (RBA):

For any fiction f , every proposition p that we believe to be true in the actual world, we are invited to imagine as part of f (i.e. p is true in f), unless p is excluded (in some sense) by another *primary* or *secondary* truth in f .

It's clear that the RA and RBA are not empirically equivalent. While the RBA would expect individuals to diverge in the propositions they import into a fiction, given that individuals hold different beliefs, the RA should expect conformity between readers over the propositions imagined, given that what is actually true is not reader-relative. Consequently, we should be able to use empirical results to differential between these assumptions.

Unfortunately, the present empirical evidence are poor adjudicators of the disagreement, for the studies concentrate upon whether readers import commonly believed propositions into the relevant fiction, and not propositions over which there is considerable disagreement. For example, very few readers would endorse the background assumption that causing one's friend

to be fired from their job would produce immense pride in oneself. Rather, almost universally, the presumption is that immense guilt would be felt. As it stands then, the data fits both the RA and RBA equally well. In order to empirically adjudicate between the assumptions, we require cases over which, while there is a fact of the matter, there is considerable disagreement between readers as to the truth.

Take the following Gernsbacher *et al.* (1992) like story:

Tom had two hungry children at home to support, and no opportunity for work. His friend Joe worked at the local 7-11 to get money to support his studies. Tom would often go to visit Joe at the store, and one night when Joe was cleaning, Tom realized the cash register was open. He quickly weighed up his options. He could opportunistically take a twenty-dollar bill without Joe noticing, and feed his family for the week, or turn his gaze to resist the temptation. Tom knew that if he took the money he would be putting Joe's job at risk, and Joe had been a good friend to him. Yet, he also knew that Joe was bright and would probably get another job. Plus, his children were very hungry. In the end, he decided to do the morally right thing.

We can presume that not everyone will share the same convictions over what Tom ought to do in this scenario. Some will be convinced he has a greater obligation to his children than Joe, and besides it's unlikely Joe's employers will connect the missing money to him. Others, however, will maintain that regardless of Tom's obligations to feed his children, he shouldn't betray the trust of his friend, and that besides there's always an honest way to make money and feed your children. Thus, according to the RBA, if as per the Gernsbacher *et al.* (1992) study, participants were provided with one of the following target sentences,

(3) Tom took the ten-dollars

(4) Tom averted his gaze

we shouldn't expect the average reading time of the sentences to differ significantly. In contrast, this is exactly what the RA would hypothesise. After all, it cannot be that both sides on this matter are equally correct. Either Tom is morally right to steal the money or not.⁶ Consequently, whichever of the target sentences reflects the morally wrong action should be read more slowly than the other. Thus, in this case at least, the RBA and RA will predict divergent results from participants' answers.

⁶ If one is not a realist about morality, and thus finds this presumption unconvincing, there are equally suitable non-moral examples, such as concerning public policy matters.

However, unfortunately, as we have mentioned, none of the studies Friend advances as evidence for the RA provide these tricky cases that split popular opinion. Only by considering such cases can we suitably discriminate between whether the empirical evidence supports the RA or one of its competitors, the RBA. This doesn't entail, however, that there are no other reasons to prefer one of these assumptions over the other. Indeed, it should be quite clear that, without strong empirical support in favour of the RA, we ought to give the RBA a far higher prior probability for theoretical reasons.

The RA requires that we import into our fictional imaginings propositions which are actually true, if the suitable question arises. Here, though, straight away is a stumbling block for the RA in its competition with the RBA. How are we as readers supposed to be aware of the suitable proposition to choose when the question arises, between p and not- p , if we do not already know which is true? That is, if we do not already *believe* the proposition! Of course, it is perfectly possible for us to imagine situations that we do not believe obtain. This, after all, is what fictions ask us to do, they require us to *imagine that*. However, the RA requires something rather different of us. It requires that, when the question arises and we must choose between some proposition and its negation, without the fiction clearly excluding either option, we ought to import whichever one is actually true into our fictional imagining. Yet, the only means through which we would have to decide between the propositions in such cases is if we already believe that one is actually true, or we seek out *to form a belief* about the matter. In other words, without prompting from the work of fiction itself, we only have our own beliefs about what is true to rely upon. We do not have some *direct access* to the real world when reading a fiction, in order to supplement our imagining of the situation.

Consequently, whenever we are asked to *imagine the real world*, this is in practice tantamount to requiring us to *imagine the world as we understand it*, for although these actions are not equivalent, the latter is the best we can do. This is why we ought to accord the RBA a greater level of prior probability, for whenever we attempt to conform to the RA we can only rely on what we already know about the world around us, and thus conform instead to the RBA.

Let us admit then that the RBA is the more *prima facie* reasonable candidate to supplement our reading of a fiction.⁷ That when imagining a fiction, it ought to be conducted against the

⁷ Actually, without strong theoretical reasons to the contrary, there are other candidate explanations for the empirical data which have far greater prior probabilities than both the RA and RBA. Namely, those assumptions only requiring fictions to *invite* us to imagine *some* of the relevant real-world propositions not excluded by the fiction. As any such hypothesis would be less empirically demanding, it would automatically be accorded a higher prior probability than the RA and RBA. Unfortunately, the question of what these theoretical reasons may be for preferring the RA or RBA over these weaker hypotheses, given that no empirical data could ever provide such reasons, is far beyond the scope of our discussion.

background of *our representation* of the actual world. To accommodate this fact, can Friend not simply adapt her own position from the RA, and accept the RBA as more plausible?

Firstly, Friend certainly does not conceive of her own position as compatible with the RBA, for she proposes that her own assumption will lead to propositions being true in a fiction which are “[not only] unknown to the author; many will remain unknown to any potential audience,” (Friend 2017: 38). This result is obviously incompatible with the RBA, which treats as the stock of background assumptions to be imported into the reader’s imaginings their own beliefs about the world. Could Friend, though, not give up this supposed positive consequence of her assumption, in order to benefit from the more plausible RBA?

It seems not, for Friend wishes her RA to provide “a comprehensive default position *for which there is no alternative*,” (Friend 2017: 34. My emphasis). That is, she wishes her assumption to pick out a privileged set of truths, those of the real world, which readers can use as the backdrop to any fiction. Yet, of course, this is a possibility which the RBA denies. There is no such privileged set of truths that we can rely upon. We can only do the best with what we have. Whereas according to Friend’s RA the propositions which a fiction *invites* us to imagine are the same for each individual, assuming we have a precise meaning of ‘exclusion’, given that the set of actually true propositions is fixed and not reader-relative,⁸ according to the RBA it is possible for the same fiction to *invite* different individuals to imagine different propositions as fictional, given that individuals have different beliefs about the real world. Thus, there is no “comprehensive default position”. We are each in our own boat, and while the fiction can lead us to *imagine replacing* certain of our planks, these will not be identical for each reader.

Consequently, in this section we have seen, firstly, that there are several complications which arise in attempting to support the RA with the empirical evidence Friend advances. Due, particularly, to the lack of clarity over what ‘exclusion’ means, the lack of direct testing of the RA against competitors with rich textual examples, and the presumption that how individuals actually read informs us somehow about how we *ought to read*. Secondly, we have seen that there is a suitable competitor to the RA, the RBA, which the available empirical data cannot discriminate between, but which we have reason to assign a higher prior probability to. Lastly, we have argued there are good reasons to think Friend cannot simply adapt her own position from the RA to the RBA without losing some of her position’s purported strengths.

⁸ Note, this does not mean everyone *does imagine* the same. As mentioned earlier, mistakes may be made, and we are *permitted* to imagine more than we are *invited* to.

4. Limitations of the Reality Assumption

So far, we have argued that the empirical evidence advanced in favour of the RA fails to support it over a competing hypothesis, the RBA, which has a higher prior probability. In this section, we go beyond concerns over the strength of evidence for the RA and consider two of its limitations: i) It confuses what we can be *invited* to imagine, by talking about what is actually true, rather than beliefs; ii) The notion of exclusion is obscure, and available interpretations lead to unsavoury consequences for the RA.

4.1. Problem One: A Prescription to Imagine what is True?

There are two distinct traditions within contemporary philosophy of fiction when specifying fictional truth. The first, traced back to Lewis (1975), defines a fiction in terms of sets of possible worlds, and thus allows us to deal with fictions just as we do scientific and philosophical possibilities, with *possible-world semantics*. No reference is made in this account to what readers are expected to imagine. Rather, in order to recognise which of the more elusive propositions are true in a particular fiction, we only need to look to those which we already admit as part of the fiction and consult the relevant possible worlds. (Which of the possible worlds, and thus which counterfactuals, are relevant will depend on the *generation principle(s)* one accepts.) According to this *metaphysical* interpretation of the *truth-in-a-fiction* problem, it is solved by ‘plugging in’ the data we have available into the gamut of possible worlds.

In contrast, the second tradition made popular by Walton (1990), emphasizes fictions not as sets of possible worlds, but constructions formed through a co-operative project between a work, a *representation*, and the audience. The truths of a fiction, often incomplete and sometimes inconsistent, are introduced by *prescriptions to imagine* by the work. In their own games of make-believe, audiences then either succeed or fail to meet these prescriptions, with their own imaginings. In order to provide these prescriptions, authors can rely on a plethora of tools, prompting the imagination in one way and that, using well-worn tropes and social background assumptions on the part of the audience. Thus, according to this *pretense* interpretation of the *truth-in-a-fiction* problem, a solution is offered by looking to what the relevant work *requires us to imagine*.

As it stands, Friend’s own RA risks being a fudge of these two distinct approaches to the problem. According to the RA, every actual truth p is true within a fiction f unless another truth of f excludes p in some sense. However, as we have seen, Friend conceives of *truth-in-a-fiction*, as Walton does, in terms of what we are invited to imagine. It should be clear, however, that

being invited to imagine what is *actually true*, without any indication of what this truth is, is an unreasonable request. Unless, that is, we conceive of the request in terms of what we as individuals *believe* is actually true.⁹ This was our motivation behind assigning the RBA a higher prior probability than the RA, for as soon as we try and meet the RA's request, the best we can do is conform to the RBA. The request to imagine what is true, as a distinct act from imagining what we *believe to be true*, without any prompting as to what is actually true, is as directionless as the request to imagine a *particular* scientific truth yet to be discovered.

The imaginative difficulty which Friend's RA places us in, and the conceptual tension within her own account, is demonstrated quite clearly by her suggestion that the RA is a suitable starting assumption when approaching a fiction, because "authors expect readers to take the real world as background," (Friend 2017: 36). There are two regards in which this is mistaken. Firstly, I cannot expect you to take some information as background unless I direct you towards it, or you already possess it. Now, as the whole function of the RA is to supplement the truths provided by a work of fiction, the former case is already ruled out. Yet, this means, without appropriate prompting, the only background information I can expect you to import into your imagining is what you are already privy to – *your beliefs about the world!* Secondly, in accordance with Walton's (1990: Chs. 1-2) *pretense interpretation* of the representational arts, the author of the work should be considered to be directing our attention to certain make-believe scenarios, relying on a host of preconceptions and socio-linguistic norms. Yet, in order to successfully achieve this, the author must not expect readers to take the real world, *whatever that turns out to be*, as the background. Rather, it must either be the case that the author expects readers to (roughly) assume the world as she knows it, facilitating successful promptings to imagine, or being aware of how her readers hold different views to her own, she may accommodate these differences, building them into her attempts to communicate her invitations to imagine. This latter case is similar to the work of a translator who takes into account not only differences in language but cultural background assumptions, in order to communicate a fictional work's invitation to imagine. In both cases, however, what the author expects a reader to assume is something she is privy to, and not the "real world" as some place-holder for whatever in the end turns out to be true.

Thus, in order to take Friend's commitment to a *pretense-like* theory of fiction seriously, and make sense of an author's invitation to imagine, we cannot conceive of the background

⁹ There is the further possibility of conceiving the request in terms of what a *particular individual*, such as an informed 'ideal reader', *believes* is actually true. Yet, such requests would constitute entirely different starting points for engaging with a fiction, which it's clear Friend wishes not to endorse; see, for example, Friend's (2017: 36-8) criticisms of the *Mutual Belief Assumption*.

assumptions within these games of imagining as the RA does. Instead, these assumptions must be propositions the audience already is privy to, such as in the RBA.

We have already observed the unsuitability of the RBA for Friend's purposes, although it possesses theoretical advantages which the RA does not. However, it is worth highlighting here a further reason Friend cannot simply transfer her allegiance to the RBA – it undercuts her arguments against a competing starting assumption, the *Mutual Belief Assumption* (MBA):

For any fiction *f*, every proposition *p* mutually believed to be true by the author's society, we are invited to imagine as part of *f* (i.e. *p* is *true in f*), unless *p* is excluded (in some sense) by another *primary* or *secondary* truth in *f*.¹⁰

Friend believes the MBA is vastly inferior to the RA as the starting point for our engagement with fictions in two respects. Firstly, “restricting ourselves to a set of beliefs leaves too many aspects of the storyworld indeterminate” (Friend 2017: 37) and, secondly, assuming the mutual beliefs of the author's audience would systematically generate incorrect fictional truths:

In works produced in some societies, it would be storified that there are ghosts and fairies, that mental illness is caused by demonic possession, or that comets are omens of catastrophe sent by the gods – even if the stories in question suggested no position on these matters... [The MBA] suggests that a society's belief system is sufficient to generate these storytruths by itself. (Friend 2017: 37-8)

The RA, we are told, suffers from neither of these concerns, for the truths it generates are neither limited by the finitude nor fallibility of our beliefs.

Once we transfer our allegiance to the RBA, however, neither argument against the MBA is convincing. Just as the MBA restricts us to the mutual beliefs of the author's community, so the RBA restricts us to our own beliefs regarding what is true. One may hope, of course, that the world of our beliefs is richer and more accurate than the mutual beliefs of past, or alternative, communities. But, this will be likely at best, and hopeful at worst. There are undoubtedly some readers at present whose appreciation of certain fictional worlds would be richer and more complete if they were to take on the mutual beliefs of another community. After all, there are those who know less now than whole scientific communities have in the past, and whose moral views are far less nuanced than those of eminent circles during the Enlightenment. Consequently, whether the RBA leads to us having a more complete and

¹⁰ A proposition is *mutually believed* if most members of the society believe it, most members of the society believe most of them believe it, and so on.

accurate picture of the world than the MBA will depend on who the reader is, and who the author's community are. Consequently, according to Friend's own criteria, the RBA cannot be wholly preferable to the MBA. Rather, their relative suitability would have to be decided on a case-by-case basis.

Thus, while Friend's RA is in tension with her view of fiction as a *prescription to imagine*, the RBA, which is not, seems unpalatable for Friend's purposes for other reasons. Namely, it blocks her general arguments against the MBA. If we are to prefer the RBA over the MBA when engaging with a fiction, this will need to be a decision taken on an individual basis. Neither the RA or RBA then should be the universal starting point for engaging with a fiction.

4.2. Problem Two: Exclusion is Obscure

Friend (2017:34) is clear that "the RA is not a mechanism for generating implied story-truths... [but rather the] starting point for specifying the inputs into any such mechanism." Yet, to suitably evaluate the RA's plausibility, it's critical we know what follows from the assumption. We cannot suitably evaluate a theory until we are aware of its (likely) consequences. If it turns out that, however we interpret the mechanism for which the RA is an input, the consequences are implausible, then this must count against its plausibility.

In order to specify when we are invited, according to the RA, to imagine real-world truths as part of a fiction, Friend appeals to the notion of a fiction *excluding* truths. We are invited to imagine these real-world truths by the fiction unless it somehow otherwise 'speaks' on these matters, including through the use of genre norms, and excludes them.

Unfortunately, as it stands the concept of *exclusion* is an unhelpful guide to when we are invited to imagine real-world truths as part of a fiction and not, for it is extremely unclear exactly how we ought to interpret such exclusion. At times, Friend talks as though what matters is *consistency* with the work:

Whatever aspects of reality remain consistent with other story-truths are themselves storified...
[and] we can draw inferences accordingly. (Friend 2017: 32)

We are, from the start, invited to represent as part of the storyworld anything that obtains in the real world that does not contradict other features of the work. (Friend 2017: 34)

Yet, on other occasions, emphasises the importance of *appropriateness* and *appreciation*:

To construct an *appropriate representation* of the storyworld, we might need to reject real-world truths for some aspects of the story. (Friend 2017: 38. My emphasis)

[G]iven the Reality Assumption, we resist abandoning the truth unless we must do so to *understand or appreciate a story*. (Friend 2017: 38. My emphasis)

It is, then, thoroughly unclear which criteria we should use when assessing whether we have been invited to imagine a real-world truth by a fiction or not. Yet, it is necessary for the sake of the RA that we show *some* interpretation of ‘exclusion’ has the possibility of providing a plausible means to demarcate which real-world truths should be imported into our imagining, and which shouldn’t. Without such an account, we lack the assurance that the RA doesn’t lead to implausible consequences.

Given Friend’s talk of *consistency*, *appropriateness*, and need for *appreciating a story*, there are two broad interpretations of ‘exclusion’ open to Friend. Firstly, in terms of some *strict* exclusion, such as logical or metaphysical consistency. Under this criterion, readers would be invited to imagine a real-world truth p as part of the fiction f as long as some other truth of f was not either logically or metaphysically inconsistent with p . Secondly, we could understand exclusion in a *weaker* sense, using the notion of *best explanation*. Under this criterion, readers are invited to imagine a real-world truth p as part of the fiction f unless p fails to form part of the best available explanation of f ’s other truths. Call the version of the RA incorporating the *strict* interpretation of exclusion the *strict* RA, and the version incorporating the *weaker* interpretation the *weak* RA. If we can show that either of these interpretations of Friend’s RA fail to entail unsavoury consequences, then we will have certain assurances over the RA’s plausibility. However, in contrast, if neither can, then we are still owed an account of what exclusion means, and assurances of RA’s plausibility.

Unfortunately, as we shall now show, it appears we are still owed such an account. For while the *strict* RA over-privileges reality, by forcing us to read fictions in a fashion that it shouldn’t, the *weak* RA fails to privilege reality enough, undermining French’s proposal. Let us begin with the *strict* RA.

Logical and metaphysical consistency are low bars to pass. There will be many real-world truths which a work, with the genre constraints and conventions it is embedded within, fails to logically or metaphysically preclude. Yet, according to the *strict* RA, a fiction’s failure to so preclude a real-world truth p is tantamount to readers being invited to imagine p as part of the fiction. Consequently, even if there are other equally fruitful, or *more fruitful*, readings of the fiction available not including p , the *strict* RA will always require we interpret the work in line

with *p*. To use Friend's (2017: 33) own example of Henry James's *The Turn of the Screw*, because ghosts do not actually exist, and it is *possible* to provide a Freudian interpretation of the work which avoids the need for supernatural entities, the *strict* RA will automatically insist on the Freudian interpretation over such supernatural interpretations, even if it is ultimately less rewarding. Thus, just as Friend (2017: 39) complains of other theories of fictionality, the *strict* RA "closes off interpretative options that should not be closed off by a philosophical account of story-truth."

Nor would such mandates on matters of interpretation be unusual. Consider the following vignette:

The Potter tribe regularly prayed to their personal god, Hedwig, sacrificing small animals in hope of a good harvest. On the occasion of such sacrifices, members of the tribe would collectively have a vision of Hedwig, providing them with detailed instructions for the following year's sacrifices. Although devout believers, due to harsh terrain, rudimentary tools, and the cunning of animals, some years the tribe were unable to supply the demanded sacrifices. While each year in which they completed their sacrifices was full of excellent harvest and wellbeing, those years in which they failed were full of hunger, thirst and death. This only increased the strength of their faith in Hedwig's powers.

Without any additional interpretative constraints due to further text, conventions or genre, it's clear we have two broad interpretations open to us here. Either we imagine that the Potter tribe's fate is in the hands of a deity, or we imagine that the tribe's visions are non-veridical, a product of some hallucinogenic, and that the harvest trends expressed in the story are a product of pure chance. Until further such constraints arise, we shouldn't be forced either way in our interpretation of the story. Yet, this is exactly what the *strict* RA requires, given that the story fails to preclude the non-existence of Hedwig. Thus, we are invited to imagine that Hedwig doesn't exist, as this is the way of the actual world, and all of the other propositions which follow from this interpretation of the vignette. This simply places too great an interpretative restriction upon us as readers. While it may turn out that the best reading of the story is that, indeed, Hedwig fails to exist, we shouldn't be *required* to read the story in this way. After all, it is very plausible, given the success of the Potters' sacrifices, that a supernatural agent does exist in the story. The *strict* RA then suffers from the same shortcoming as the *Reality Principle*, for it over-privileges the actual world by forcing readers to take interpretative steps which no philosophical theory should.

Does the *weak* RA serve readers any better? At first glance, yes. By only requiring that we import into our imaginings real-world truths when they form the *best explanation* of a fiction's

other truths, the *weak* RA succeeds in not forcing readers to take interpretative steps which they shouldn't. However, with this interpretative leniency comes a price, for it's unclear that the *weak* RA privileges real-world truths to the extent Friend requires.

The RA requires “we default to the assumption that familiar behaviour can be explained in the ordinary way, and we do not give up that assumption unless required to do so,” (Friend 2017: 34). Thus, the purpose of the RA's exclusion constraint is to explain when exactly *the presumption of reality* is defeated. Yet, according to the *weak* RA, we are only invited to imagine a real-world truth as true within a fiction *f* if it forms part of the *best explanation* of *f*'s other fictional truths. Yet, this would make real-world truths no different to other competing possible auxiliary assumptions. Just as we should introduce other propositions into our reading if they help explain elements of a story, and leave them out if not, so the *weak* RA dictates we are only *required* to include real-world truths in our imagining of a fiction if they form part of the best overall interpretation of the work. Thus, the *weak* RA, rather than treating the inclusion of real-world truths within fictional works as the status quo, conceives of real-world truths as just some in a whole host of possible background assumptions we could use in our imaginings of fictions. Gone is the presumption of including real-world truths in fictions. Instead, we should just import those propositions which best explain a fiction's other truths, regardless of whether they are real-world truths or not. To this extent, the *weak* RA is wholly unsuitable for Friend's purposes, for it fails to respect the “bias in favour of reality” which Friend (2017: 34) expects the RA to articulate.

Neither interpretation of ‘exclusion’ which can be drawn from Friend's own discussion of the RA then seems suitable for her purposes. While the *strict* interpretation places too great a constraint on our interpretation of works of fiction, the *weak* interpretation fails to respect the RA's intended reality bias. As it stands, unless Friend can provide us with another reasonable interpretation of ‘exclusion’, we must admit that either the RA is implausible for reasons similar to the *Reality Principle*, or contrary to initial impressions and Friend's intentions, the RA fails to privilege the real-world. Neither option is palatable for Friend, we are owed another account.

5. Conclusion

In proposing the RA, Friend has attempted to provide insight into how the real world permeates into fictions and informs our understanding of them, without presenting a complete account of the truth-conditions of fictions. In taking on this more restricted task, she has hoped to circumvent the failings of the *Reality Principle* and other *generation principles*. However,

as this paper has shown, the RA itself faces considerable challenges. While the empirical evidence that Friend advances in favour of the RA fails to support it over a weaker assumption, the RBA, which we have reason to assign a higher prior probability to, the RA itself suffers from theoretical weaknesses. Namely, that the RA is incompatible with Friend's own commitment to a *pretence* theory of fiction, and neither of the available interpretations of 'exclusion' within the RA are suitable for Friend's purposes.

As the failings of previous *generation principles* have shown us, the practice of the reading of fiction, and the inferring of fictional truths, is a multifarious affair. The same challenges that face these principles ultimately face the RA – it is difficult to extract from the practice of reading a general principle, or assumption, which suits all works of fiction and reading circumstances. There should be little doubt that our interactions with the actual world play a substantial role in our interactions with fictions. However, this paper has called into question that we should always treat the actual world as our starting point when reading works of fiction.

References

- Currie, G. (1990). *The Nature of Fiction*. Cambridge: Cambridge University Press.
- Friend, S. (2017). The Real Foundation of Fictional Worlds. *Australasian Journal of Philosophy* 95(1): 29-42.
- Gernsbacher, M. A., H. H. Goldsmith, and R. R. Robertson (1992). Do Readers Mentally Represent Characters' Emotional States? *Cognition and Emotion* 6(2): 89-111.
- Gernsbacher, M. A., B. M. Hallada, and R. R. Robertson (1998). How Automatically Do Readers Infer Fictional Characters' Emotional States? *Scientific Studies of Reading* 2(3): 271-300.
- Grice, H. P. (1975). Logic and Conversation. In: P. Cole and J. L. Morgan (eds.), *Syntax and Semantics*. Vol. 3. New York: Academic Press, 41-58.
- Haenggi, D., M. A. Gernsbacher, and C. A. Bolliger (1994). Individual Differences in Situation-based Inferencing during Narrative Text Comprehension. In: H. van Oostendorp and R. A. Zwaan (eds.), *Naturalistic Text Comprehension*. Norwood, NJ: Ablex, 79-96.
- Lewis, D. (1978). Truth in Fiction. *American Journal of Philosophy* 15(1): 37-46.
- Walton, K. L. (1990). *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, MA: Harvard University Press.
- Weisberg, D. S. and J. Goodstein (2009). What Belongs in a Fictional World? *Journal of Cognition and Culture* 9(1): 69-78.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
- Woodward, R. (2011). Truth in Fiction. *Philosophy Compass* 6(3): 158-67.