**"Because I Said So: Authorial Authority and Truth in Fiction"**

Jonathan Ichikawa

Draft: February 14, 2006

Work in progress. Do not cite. Comments welcome.

ichikawa@brown.edu

**Abstract.** I question alleged limitations on the author's ability to make propositions true in the fiction. I defend a strong principle of authorial authority according to which authors can make whatever they want true in the fiction. I defend this claim from Richard Hanley's recent argument that it is very difficult to make certain propositions true in fictions, and from the standard view in philosophy of fiction according to which the "puzzle of imaginative resistance" poses a threat to authorial authority. I close with an attempt to explain why we seem to have contrary intuitions.

For any fiction, certain propositions or sentences are **true in the fiction**. Others aren't. These representations include the things that the fiction literally tells us – but they also include things that are not made explicit. It is true in the fiction that Harry performs magic – this is explicit – but it is also true in the fiction that Harry has kidneys; this fictional truth is arrived at less directly, presumably by reference to things we actually know about actual people. Although authors don't explicitly lay out these latter fictional truths, they do have the authority to include or exclude them. It's true in the fiction that Harry has kidneys, but if the author wanted it not to be, the insertion of a single line would do it.

So authors have considerable control over what is true in the fiction. Call this a principle of **authorial authority** for truth in fiction. My project concerns the limitations of authorial authority. Although authors have *considerable* control over truth in fiction, many philosophers have argued that they do not have *absolute* control. Certain things, they say, cannot be true in the fiction. I attempt to refute these alleged limitations and defend the **strong principle of authorial authority**:

> (SAA) For any fiction, for any proposition, the author has the authority to make that proposition true in that fiction.[1]

To defend this principle is to refute challenges to it.

**Impossibilities True in the Fiction.** Richard Hanley (2004) has a recent defense of a version of Lewis's possible-worlds analysis of truth in fiction. In its simplest version, Lewis's analysis says that what's true in the fiction is what is true in the nearest possible worlds in which the story is told as known fact. Hanley goes to some length to respond

[1] I pass over a complication about individuating fictions; I assume for brevity that the fictional truths are not essential to a fiction. This is dispensable for my project; I trust my intent is clear.

to the objection that some fictions are impossible – true in no possible world. He argues that it’s not nearly so easy as one might think to have central contradictions true in fictions.[2]

The first ‘argument from impossible fictions’ Hanley cites is presented by Gregory Currie.[3] Currie suggests that a fiction about a person who refutes Gödel’s Incompleteness Theorem would provide a counterexample to Lewis, since there are no possible worlds in which that theorem could be refuted. Here is Hanley’s response:

> But this employment of thought experiment, telling a short *fiction about a* non-existent fiction (‘suppose someone writes a story about a Gödel-refuter’), is bound to fail. I grant that if it’s possible for it to be fictional that a contradictory proposition is fictional, then it *might* by Lewis’s own analyses be possible for a contradictory proposition to be fictional. But to insist on this interpretation of Currie’s story-about-a-story, is to beg the question. Lewis can simply deny the interpretation. The only way to avoid this fallacy is to present an *extant* fiction.[4]

This response is puzzling; it is not difficult to write such a fiction. I can do it:

> **Benny.**
> Once upon a time, there was a mathematician named Benny. Benny was struggling to make a name for himself in academia, until one day, when he discovered a flaw in Gödel’s proof of the incompleteness of arithmetic. When he first discovered the proof, he could scarcely believe his eyes – how could so many logicians and mathematicians have failed to notice the mistake? Once he diagnosed their error, though, he gained an unprecedented understanding into math and logic. Indeed, shortly thereafter, on top of his rejection of Gödel’s proof, Benny produced a proof of the *completeness* of arithmetic! Benny was the best mathematician ever. The end.

---

[2] Hanley is really concerned with fictional impossibilities that play a central role in the fictions; he admits that there are fictions in which inconsistent statements are true, which can be handled by Lewis’s “method of union.” But when inconsistencies are central to the story – that is, when the fiction cannot be divided into consistent parts – the method of union fails. These are the fictions that Hanley denies.

[3] Currie (1990), p. 69

[4] Hanley, p. 119

Hanley won't be satisfied by this response, even though it would involve referencing an actual, instead of a merely hypothetical, fiction. Hanley will deny that it is true in *Benny* that Benny refuted Gödel. He says:

> Even if a story contains the words, 'Jack refuted Gödel', that is a very long way from establishing that it is true in the story that Jack refuted Gödel. For any attempt to make it so, it may be that the best interpretation of the story is that it has an (in this respect) unreliable narrator, one who thinks that Jack has refuted Gödel, but is mistaken. (It may be that Jack has indeed made a major mathematical discovery, but thanks to the narrator's unreliability, it is indeterminate which discovery it is.)[5]

But there are serious problems with taking this line. For one thing, we can add stipulations to ensure that our protagonist has refuted Gödel, rather than accomplished some other task: *At first, people thought that Benny had made a major mathematical discovery, but that he falsely took it to be a disproof of Gödel. But after close scrutiny, they determined that it really was a disproof of Gödel; Benny proved that all true sentences of arithmetic are provable from the Peano axioms.* To suggest that *really*, Benny did something else, given this text, is just to refuse to take the text seriously.

Another problem with Hanley's suggestion is that just doesn't square with Hanley's own positive view, which is based heavily on Lewis. On Hanley's preferred interpretation, we should understand fictions "in terms of worlds where the teller really does what the actual teller only pretends to do."[6] The idea is that the author is pretending to say things he knows; what's true in the fiction is what would be true if he really were saying things he knows. But once we have Hanley's own principle in hand, application to the case in question is simple: if the teller were really doing what the author is pretending to be doing, then he'd be telling us something that he knows – namely, that

---

[5] *ibid.,* p. 12

Benny refuted Gödel. But if that's something that he *knows*, then it's true. So by Hanley's own lights, it is true in my fiction that Benny refutes Gödel. This is indeed an implication of *every* theory of truth in fiction on the table that has been put forward. (I'll give a very brief survey of views near the end of this paper.) I conclude that there is a strong presumption in favor of the possibility of fictions with central impossibilities as contents.

**'555' Telephone Numbers.** Hanley also claims that it's not as easy as we might think to make even possible things true in a fiction. His primary example has to do with the convention in fictions to use '555' as a prefix for telephone numbers. The argument is pretty simple. Hanley says that it is not plausible to suppose that it is true in many fictions that many phone numbers begin with '555'.

> If a standard mention of a '555' number generated the fictional truth that a particular number started '555', then some stories turn out to be much less ordinary than they in fact are. Consider any fiction which contains several standard mentions of '555' numbers. ... What explains the constancy in the first three digits? There seem to be two basic possibilities. First, it may be that although in the world of the fiction there is the usual divergence in these digits, for some coincidental reason, every number mentioned in the fiction has the same ones. Or, second, it may be that virtually every number in the U.S.A. begins with '555', hence it is no coincidence that the ones mentioned do. But both explanations are implausible.[7]

I agree with Hanley when he says that "the '555' number, in a standard mention, is a conventional device for communicating the (fictional) fact that someone reveals a phone number, without stipulating what that number is."[8] But I disagree that this demonstrates an interesting and substantive limitation to authorial authority. Hanley says that this

---

[6] *ibid.*, p. 114
[7] *ibid*, p. 121
[8] p. 122

consideration demonstrates that the author cannot make it fictionally true that a phone number begins with a 555 prefix just by saying so. "The author's hands are effectively tied by the convention."[9] I think Hanley is very wrong about this latter point. The reason those phone numbers don't fictionally start with '555' is that the author hasn't told us that they do. It's like speaking non-literally – if I utter, "it's raining cats and dogs," I don't imply that household pets are falling from rainclouds. Can an author make it fictionally true that a number begins with '555'? Of course he can, just as he can make it true that cats and dogs fall from rainclouds. It's just slightly trickier, because he has to make it clear that he's not uttering the words he utters in the way that is often used to invoke mere conventions. Here's a fictional passage that does the trick.

> **Jack and Jill.**
> Jack smiled. "So, can I call you later?"
> "Sure," said Jill. "My number is 555-2810."
> "Very funny," Jack retorted. "If you don't want to give me your number, that's fine. Just tell me so; don't give me a fake number. I'm not woefully ignorant of fictional conventions, you know. Obviously that's not a real number."
> Jill looked concerned. "No, honest, this is my number. I know it sounds funny, because ordinarily we only encounter '555' numbers when they're being presented in fictions as a conventional way to communicate the (fictional) fact that someone reveals a phone number, without stipulating what that number is. But for some crazy reason, the phone company actually gave me a '555' number. Here, if you don't believe me, call it right now."
> Jack wasn't sure what to think, so he dialed the number, 5-5-5-2-8-1-0 on the phone that was sitting on the table next to them. Jill's phone rang. Relieved, Jack smiled and left. The end.

It's obviously true in the fiction that Jill's number begins with '555'. So authors can make such things true in the fiction, without all that much work. They just have to make it clear that they're speaking literally in uttering words that are often used non-literally.

---

[9] *Ibid.*

(It is similarly *slightly* difficult to make it true in the fiction that cats and dogs are falling from the sky. "It was raining cats and dogs" doesn't quite cut it.)

**Imaginative Resistance.** An explicit challenge to authorial authority is the *puzzle of imaginative resistance*. It is my view that contrary to the current mainstream views on imaginative resistance, the phenomenon in question does not threaten the strong authorial authority thesis. It's generally agreed that the puzzle of imaginative resistance dates back to David Hume. In 1757, Hume wrote what is now a notorious passage in *On the Standard of Taste*, contrasting our reactions to false factual claims in works of literature with our reactions to false moral claims.[10] (The puzzle has since been generalized well beyond the moral – Walton and Yablo were influential here.) Hume's observation, which was directly about not fictions but histories, went largely unnoticed until 1994, when Richard Moran brought it to bear on emotional engagement with fiction.[11] Moran is also, as far as I know, the first person to use the term "imaginative resistance" to refer to this puzzle – or, as we shall soon see, this class of puzzles. In his example, we experience *resistance* toward the suggestion that it's ok to murder one's guest – even in the fiction.

Brian Weatherson's (2004) paper holds the key to defending the authorial authority principle from the puzzle of imaginative resistance, although it's clear that Weatherson didn't think of it as such. Weatherson observes that the puzzle of "imaginative resistance" is not one puzzle but several – at least four. The two puzzles that are of interest here are the **imaginative** puzzle and the **alethic** puzzle. The imagination puzzle is: why don't we imagine certain sentences that are given in the text of the fiction, given

[10] Hume 1757, p. 247.

that we usually imagine what the author tells us to? The alethic puzzle is: why don't certain sentences in the text of the fiction come out true in the fiction, given that the author usually makes sentences true in the fiction just by writing them?

Once this distinction is clear, the way is cleared for my defense of the strong principle of authorial authority, along with a contribution to the imaginative resistance debate: there is no alethic puzzle. The author *can* make whatever he wants true in the fiction just by saying so. There is a phenomenon of imaginative resistance, but it's not about truth in fiction; it's about our psychological responses to fictions. The only thing that needs explaining is the imaginative puzzle.

The first thing to recognize is that the two puzzles are indeed separable. In Moran's presentation of the puzzle, he says: "If the story tells us that Duncan was *not* in fact murdered on Macbeth's orders, then *that* is what we accept and imagine as fictionally true." But it's one thing to accept something as fictionally true, and another to imagine it. We accept it as fictionally true that Sherlock Holmes had a mother, even if we do not imagine it. And we may imagine fictional scenarios which we know not to be fictionally true, as when we consider how Harry might have turned out if his parents hadn't died. It may be that there is some close connection between truth in fiction and imagination, but it is not identity. (This must be so – one is a claim about the fiction, and one is a claim about some reader or readers.)

Let's work with an example of a fiction that generates resistance:

> **Filicide.**
> Bob killed his baby daughter because he thought it'd be funny to make her suffer. His neighbors thought he'd done something morally wrong, but

[11] Moran 1994, p. 95.

> they were mistaken; killing one's baby daughter is morally permissible, as long as it's funny.

It's true in the fiction that Bob killed his daughter, just because I say so, and if you engaged in the fiction, you imagined that Bob killed his daughter, just because I instructed you to. The first sentence doesn't generate resistance. But the moral claim does. If you are a typical reader, you reject the fictional claim about the permissibility of funny filicide. You do not – whether because you refuse, or because you cannot comply – imagine that it's ok to kill one's daughter. And you may also say, "it's not even true *in the fiction* that it's ok to kill one's daughter!" These are two separate resistances, once we see that your imagining is not the same as the proposition's being fictionally true.

But once we separate the resistances, I don't see any compelling reason to think that there is any truth in fiction puzzle; I think explaining the imaginative version is all there is to do. If you go on to say that it's not even true *in the fiction* that it's ok, you are either hyperbolizing or you're mistaken. Why? Well, to start, note the following asymmetry between the imagination and alethic puzzles:

> **Imagination puzzle**: the reader does not imagine that funny fillicide is ok, even though the fiction tells her to. This provides an exception to the general principle, "when the fiction says p, the reader imagines that p." That's puzzling.

> **Alethic puzzle**: the reader does not believe that it's true in the fiction that funny fillicide is ok, even though the fiction says it is. This provides an exception to the general principle, "when the fiction says p, it's true in the fiction that p." That's puzzling.

The latter would be prima facie puzzling if it were true, but it's not. The fact that a reader doesn't *think* something's true in the fiction doesn't provide a counterexample; there's a gap between our beliefs about of truth in fiction and truth in fiction – a much

more substantial gap than there is between our judgments about what we're imagining and what we're actually imagining. So we don't have an obvious exception to the truth in fiction principle, the way we do to the imagination principle.

One might respond, we're generally pretty good at recognizing what's true in the fiction, and so our belief that it's not true in the fiction is *good evidence* that it's not true in the fiction. I certainly think we're at least *pretty* good at recognizing fictional truths, but this doesn't settle the matter. This is especially salient in light of the fact that most people – including most of the people who are reporting these intuitions about truth in fiction – are not typically reflective about truth in fiction. It may be that our judgments in this matter are as yet unrefined. Such intuitions are particularly likely to be unreliable in deviant or unusual cases – as the cases which generate imaginative resistance surely are. So we should at least be open to the possibility that we're going wrong when we deny that it's true in the fiction that funny filicide is morally permissible – especially if there are good reasons to accept the strong authorial authority thesis, and we have ways of explaining away the intuition that it's not true in the fiction that filicide is ok. Establishing these two claims is the remainder of my project.

I think there are three positive reasons to think that it is true in *Filicide* that funny filicide is morally permissible. The first reason is simple theoretical elegance. Once we see that there are separate puzzles, if we can explain everything funny that seems to be going on in terms of the imaginative puzzle plus other facts about readers' psychology, we have no reason to suppose that we have something funny going on at the non-reader-psychology level of truth in fiction; we can comply with the general principle I began with: *in general, if the author says that p, then p is true in the fiction.* If we can tell a

plausible story that maintains this feature, this is a desirable consequence of the view. Obviously, much rests on the plausibility of that story, which I will attempt to demonstrate in what follows

The second reason I have in mind has to do with morally repugnant claims that are uncontroversially true in fictions. The idea is that there are some fictions which generate the same sort of moral disgust that the Bob filicide story generates, but where it is clear that the morally repugnant claims *are* true in the fiction, even if we wouldn't like to admit it or describe it in those terms. There could be a fiction in which a scientist discovers that black people are always stupider and more violent than white people. There could be a fiction about a white supremacist who fearlessly defies considerable social pressure and discovers that there was no Holocaust, and that the public has been massively misled by a giant conspiracy of the world's Jews. These would be terrible fictions, and we should not read them, and we should abhor those who would write them. But there is very little doubt that there could be fictions like this.

If there were fictions like this, and we were asked whether it's true in the fiction that the Jews of the world have duped everyone into believing in the Holocaust, perhaps the all-things-considered appropriate response would be to refuse to answer the question: "how can you even think such a thing," we might say. Or: "that fiction isn't even worthy of our consideration." Nevertheless, it seems clear that the literal answer to the question is yes: it *is* true in the fiction that there was no Holocaust, as much as it pains us to have to think about it and say it. (Indeed, it seems that it's exactly that these repugnant claims are true in the fiction, in virtue of which the fiction is so repulsive.

Why else would we object to them so?) I submit that this parallels the experience of *Filicide*; there, as here, we should admit that the repugnant claim is true in the fiction.

The third reason I think we should admit that it's true in the fiction that funny filicide is ok is that this follows from every plausible theory of truth in fiction that has been proposed. Here's a quick survey:

**David Lewis: counterfactual:** *p is true in the fiction if and only if, were the story told as known fact instead of fiction, p would be true.*

This is pretty straightforward – if the story I tell above were a story that I knew, instead of a story I made up for a philosophy of fiction paper, then Bob's action would have been permissible. This is obvious: if I said that it was permissible and knew that it was permissible, then it was permissible.[12]

**Kendall Walton: prescription to imagine:** *p is true in the fiction iff the fiction generates a prescription to imagine p.*

There's clearly a prescription to imagine that Bob's action was permissible; that's why imaginative *resistance* occurs. We recognize that we've been instructed to imagine that p, and that's why it's interesting that we don't do it.

**Gregory Currie: reasonable belief attribution:** *p is true in the fiction iff it would be reasonable to believe of someone who sincerely asserted the story that he believed p.*

---

[12] I here assume that we can evaluate counterfactuals with necessarily false antecedents, and that they are not trivially true. This is not controversial. Lewis himself thinks that counterfactuals of the form *if [something impossible] were to be the case, then p would be the case* are all trivially true. But this is implausible. Here is a counterexample: *if I were a female bachelor, I would be married*. This counterfactual has a necessarily false antecedent, because all bachelors are male, yet the entire counterfactual is false; I would not be married if I were a female bachelor. It is also worth mentioning that even if the view in question is right, and all such counterfactuals are trivially true, then the repugnant moral claim still turns out true in the fiction – along with everything else.

Some people have crazy moral beliefs, and sometimes it's appropriate to attribute crazy moral beliefs to people. If confronted with someone who sincerely reported my story, it would be perfectly reasonable to attribute the belief that funny filicide is permissible. If sincerely saying that funny filicide is ok isn't good evidence that one believes it's ok, then I don't know what is.

So far, I've argued that there's room for fallibility in our attributions of truth in fiction, and that there's good reason, independent of our intuitions about resistance cases, to suppose that the resisted statements *are* true in the fiction. To close the coffin on the alethic puzzle, the last remaining step is to offer some particular reasons to think that our intuitions would go wrong in these cases. Thus far, we have a tension: the intuitions, on one side, that it's not true in the fiction that funny filicide is ok, and the theories, on the other, which suggest that this claim is true in the fiction. In this section, I will have a few things to say about why we should trust the theory over the intuition in this particular case.

**Tangled Puzzles.** The first argument is simple. It also seems to me to be one of the most compelling. We say there's an alethic puzzle in addition to an imaginative puzzle, but observe how easy it is to confuse truth in fiction with the willingness and ability to imagine: indeed, before Weatherson's paper in 2004, everyone who had written on imaginative resistance seems to have been making this error. This is good evidence that our pre-theoretic intuitions are not fine-grained with respect to the distinction between the alethic and imaginative puzzles.

**Truth in Fiction, Truth about Fiction.** Here is a second argument: the grammar of the 'truth in fiction' locution may mislead us. The datum that needs explaining is our unwillingness to say:

(1) It's true in the fiction that Bob's action was morally permissible.

I suggest that our unwillingness here may derive from a proper unwillingness to say:

(2) It's true that Bob's fictional action was morally permissible.

We read the fiction and we say, "no! It's not true that Bob's action was permissible. The author has no special authority to say *that*!"[13] The reader is right if he means to say that the author has no special authority to *pass moral judgment* on Bob – if he means to be denying (2). It is false – *literally* false – false *in the actual world* – that Bob's action was right.[14] But of course, the author *does* get special authority to tell us whether it's true *in the fiction* that the action is right. One reason we might reject (1) is that we confuse it with (or wrongly think it entails) (2), a claim about the actual world. The idea is that we take the author to be claiming something about the actual world. This is an interesting psychological fact about us readers, but since fictional truth doesn't really entail an actual truth in this way, there's no reason here to think that it can't be fictionally true that Bob's action is permissible. The key is to see that (1) does not imply (2); so we need not reject (1) in order to reject (2).

**Truth in Fiction as a Kind of Truth.** Another reason we may want to reject the fictional truth claim is that we wrongly think that fictional truth is a kind of truth. We say, "funny fillicide is not permissible anywhere – not even in fictions!" The idea here is

[13] C.f. the discussion of 'special authority' in Gendler (2000) and Weatherson (2004).

[14] I assume here that we can make sense of fictional actions having real properties like rightness and wrongness. Things get more complicated otherwise, but I think we can still tell the story.

that we're thinking that if something is true in a fiction, then it's *sort of* true, or *possibly true*, or *true somewhere*, or something else that weakens our commitment to considering the claim to be absolutely false.

But that's not how truth in fiction works – I observed above that we can have contradictions true in fictions. We also often have indeterminacy of truth in fiction. "Truth in fiction" is in some respects an unfortunate locution, because truth in fiction doesn't really behave very much like truth. This was one of the motivations for Gregory Currie's view about reasonable belief attribution: fictional truth makes more sense modeled on belief than modeled on truth. The analogy is instructive – we may describe someone's belief set – the totality of a person's beliefs – on the "true in a world" model. That is, we may speak of the *belief world* of an individual. "p is true in S's belief world" can mean "S believes that p". We may now say that pretty much anything may be "true in a belief world", without thereby undermining our commitment that certain statements never be true anywhere or in any way. If we understand truth in fiction on this model, then we should not resist allowing morally repugnant claims to be true in fictions, any more than we should resist allowing them to be true in belief worlds.

## References

Currie, G. (1990): *The Nature of Fiction*. Cambridge University Press

Currie, G., and Ravenscroft, I. (2002): *Recreative Minds*, xxx

Gendler, T. (2000): "The Puzzle of Imaginative Resistance", *The Journal of Philosophy*, 2: 55-81

Gendler, T. (forthcoming): "Imaginative Resistance Revisited", in *The Architecture of the Imagination*, Nichols, ed., Oxford University Press (forthcoming)

Hanley, R. (2004): "As Good as it Gets: Lewis on Truth in Fiction", *Australasian Journal of Philosophy*, 82 (2004), pp. 112-128, March 2004

Hume, D. (1965): "Of the Standard of Taste", in *Of the Standard of Taste and Other Essays*, Hume. Indianapolis, Bobbs-Merrill: 3-24

Lewis, D. (1978): "Truth in Fiction", reprinted in Lewis, D. (1983), *Philosophical Papers, Vol. I.*, Oxford University Press pp. 261-80

Matravers, D. (2003): "Fictional Assent and the (so-called) 'Puzzle of Imaginative Resistance'", in *Imagination, Philosophy, and the Arts*, Kieran and Lopes, eds., London, Routledge: 91-106

Moran, R. (1994): "The Expression of Feeling in Imagination", *Philosophical Review* 103(1)

Nichols, S. (2004): "Imagining and Believing: The Promise of a Single Code", *Journal of Aesthetics and Art Criticism*, Special issue on Art, Mind, and Cognitive Science, 62, 129-139

Walton, K. (1990): *Mimesis as Make-Believe*. Harvard University Press

Walton, K. (1994): "Morals in Fiction and Fictional Morality", *Proceedings of the Aristotelian Society*, Supplementary Volume 68

Walton, K. (forthcoming): "On the (So-Called) Puzzle of Imaginative Resistance", in *The Architecture of the Imagination*, Nichols, ed., Oxford University Press (forthcoming)

Weatherson, B. (2004): "Morality, Fiction and Possibility", *Philosophers' Imprint* 4(3).

Yablo, S. (2001): "Coulda, Woulda, Shoulda", in *Conceivability and Possibility*, Gendler & Hawthorne, eds., Oxford University Press pp. 441-492