1 Philosophical fictionalism

The project that we pursue in this book involves the application of a position we call *philosophical fictionalism*, in order to deal with a variety of philosophical puzzles or problems. Philosophical fictionalism is a particular species of a current movement in philosophy that marches under the banner 'fictionalism'.[1] Recently, there has been a surge of interest in philosophical views collected under this banner, but in much of the literature, the relationships among various ideas connected to fictionalism have been muddled, resulting in a fair degree of confusion. Thus, despite the increase in the production and discussion of theories classified as *fictionalist*, there is still a need for further clarification of the classification itself. For a start, it is important to recognize that what philosophers call "fictionalism" is really a genus of theorizing that we should analyze further into species and varieties. Our aim in this first chapter is to set out some of the details of different species of fictionalism, locating philosophical fictionalism within the genus and motivating our preference for it.

1.1 Two species of fictionalism

We start with what we take to be an intuitive minimal requirement that an instance of philosophical theorizing must satisfy in order to count as a case of fictionalism: It must make some appeal to the notion of fiction. While we think this is a fairly obvious condition for being properly characterized as a "fictionalist account", it bears noting that endorsing just this already puts us at odds with how a number of philosophers have thought about fictionalism in philosophy. For example, quite a few recent articles and books classify Bas van Fraassen's constructive empiricism as a *fictionalist* view,[2] despite the fact that

[1] While we are treating 'philosophical fictionalism' as a term of art for a *species* of fictionalism, we should note that at least some authors (for example, Daly [2008]) have also used it to designate the genus of fictionalism as a whole.

[2] Cf. Field (1980), Yablo (2001), Hussein (2004), Kalderon (2005), Rosen (2005), Blackburn (2005), Nolan et al. (2005), Sainsbury (2010), and Eklund (2011), for a start.

van Fraassen makes no appeal to the notion of fiction and even lists fictionalism as one of the untenable positions regarding science.[3] This is not to say that one cannot *construct* a fictionalist view on the basis of what van Fraassen says about science. But that is different from characterizing van Fraassen's view as fictionalist. It also seems possible to *recast* van Fraassen's view "in a fictional spirit", as Gideon Rosen does.[4] But what Rosen does seems more an instance of what we might describe as *extra*-fictionalism about a "not-already-fictionalist" theory rather than a proper identification of the theory's status on its own.[5] Of course, Rosen and others are free to use 'fictionalist' in order to describe van Fraassen's view. But doing so seems to rid the term of much significance, since many accounts that we would not wish to characterize as fictionalist can be seen at least in some ways as being "in a fictional spirit" (for example, expressivism in ethics, since it maintains that ethical claims do not express beliefs but, instead, express some other attitude).

Granting our intuitive requirement, it is useful to note a contrast between two ways one can make an appeal to the notion of fiction in one's theorizing. The first is to bring in the notion in a comparative manner for the purpose of analogizing how one is treating (or proposes treating) the target of one's theorizing. Theorists often indicate that this is what they are doing when they offer a typically vague instruction to treat the subject of theorizing "as a kind of fiction".[6] For what we call *comparative fictionalists*, the appeal to fiction is *external* to the account itself and functions more as a suggestive guide. The proposed treatment of the target of theorizing need not itself explicitly appeal to the notion of fiction at all.

In cases of comparative fictionalism, theorists usually have in mind certain claims they want to make about some topic (or, if they are being more careful, about a certain fragment of discourse, putatively about that topic – more on this point in section 1.2) and they highlight various similarities between what they want to claim about their target of theorizing and what supposedly holds for works of fiction. The similarities noted can be from a wide variety of categories. Often, they have to do with the semantics or pragmatics of claims from the relevant discourse (for example, the claims are not true,[7] or the claims are not asserted but are only "quasi-asserted"[8]). In other cases they are epistemological

[3] van Fraassen (1980), p. 2. [4] Rosen (1994).

[5] This is also how we see David Lewis's (2005) description of Simon Blackburn's quasi-realism as an instance of fictionalism. Blackburn (2005) explicitly distinguishes his view from fictionalism (and cites a much earlier article where he already contrasts his view from what he calls an "as if" view). We think that is sufficient to make quasi-realism not a fictionalist view, but here too one could construct a fictionalist view on its basis or give an *extra*-fictionalist recasting of it (as Lewis does).

[6] Rosen (1990), Yablo (2001), Daly (2008), and Eklund (2011) all attribute this instruction to fictionalism.

[7] Field (1980), Nolan et al. (2005), and Balaguer (2009).

[8] Yablo (2001), Daly (2008), Kroon (2011).

(for example, the contents of the claims of the relevant discourse are only to be accepted rather than believed[9]) or metaphysical (for example, the putative entities described in the claims of the relevant discourse do not exist or are not real[10]), but in any case, the similarities are held to underwrite an analogy between (discourse about) the target of theorizing and fiction. Most of the accounts that have been offered as instances of fictionalism in the philosophical literature are cases of comparative fictionalism.

The second way that one can appeal to the notion of fiction in one's philosophical theorizing is by making explicit use of it in the details of one's account of the subject of theorizing.[11] In such cases, the appeal to fiction is *internal* to the account itself. This more integrated use of the notion of fiction is part of what we consider to be distinctive of the species we call *philosophical fictionalism*. The notion of fiction has a more direct role in the theorizing involved in instances of philosophical fictionalism. In cases of comparative fictionalism, the theorizing about the topic need not itself make use of anything that counts as fiction. In contrast, while an instance of philosophical fictionalism applies the notion of fiction in the account offered, it is not necessary that the subject of theorizing be similar to fiction in any way that underwrites an analogy.[12]

To explain further the integral role that fiction plays in cases of philosophical fictionalism, it is helpful to analogize the relationship between philosophical fictionalism and fiction with one way we might understand the relationship between philosophical logic and logic. Philosophical logic appeals to and applies elements of logic (that is, techniques and devices from logic) in order to resolve philosophical issues that do not have to do with logic (at least directly). A familiar example is Bertrand Russell's rejection of the subject-predicate analysis of sentences involving definite descriptions (and common names), in favor of a quantificational analysis, to resolve certain linguistic (and metaphysical) puzzles.[13] Similarly, philosophical fictionalism applies the

[9] Vaihinger (1911[1924]), Putnam (1971), Neumann (1978), Kalderon (2005 and 2008).
[10] Putnam (1971), Woods and Rosales (2010), and Kroon (2011).
[11] Of course, for one who makes explicit use of the notion of fiction in the details of his account, a comparison with fiction can still be part of the appeal of doing so. The key difference is that comparative fictionalism just makes the comparison and leaves it at that. It bears noting that one might make the comparison and then go on to integrate fiction into the account of what to do about, for example, the semantic infelicity of the claims from a fragment of discourse. Such a theorist then switches over from comparative fictionalism to philosophical fictionalism. One might view Field (1989) (in contrast to Field [1980]) and Rosen (1990) as coming close to this.
[12] This point undercuts Caddick Bourne's (2013) worry that even if one makes an explicit appeal to fiction in an operator that one postulates to be at work in some discourse, that still might not make the semantics of the discourse the semantics of fiction (p. 153). We agree that it might not, but even dissimilarity from works of fiction on this front (and others) would not automatically block a view from counting as philosophical fictionalism and thus fictionalism.
[13] Russell (1905). Cf. Grayling (2001).

notion of fiction and makes use of elements of certain kinds of fiction to resolve philosophical problems or puzzles that do not necessarily have to do with fiction. As a preliminary example, consider Gareth Evans's account of existential claims (especially negative existentials), which explains the linguistic functioning of this fragment of discourse in terms of the operation of a game of make-believe.[14]

In contrast with philosophical logic, the philosophy *of* logic is concerned with issues that arise in, or for, a logic – for example, about how to understand the proper treatment of conditionals in logic or about the notion of logical consequence. And, just as it is important to heed the difference between philosophical logic and the philosophy of logic, it is important to do the same with respect to a parallel distinction between philosophical fictionalism and the philosophy *of* fiction. The latter investigates philosophical issues and concepts that pertain to fiction itself – for example, about the putative ontological status of so-called "fictional entities", the truth-values and semantics of sentences that appear in works of fiction (what we, following others,[15] will call *fictive discourse*), and the semantics of claims about works of fiction and the putative entities and situations they portray (that is, *metafictive* and *transfictive* discourse).

It is crucial to be clear about the distinction between philosophy of fiction and fictionalism in order to avoid two kinds of problems. The first is mistaking work done in one arena as automatically applicable in, or extendible to, the other. The most common version of this error is thinking that applying work done in philosophy of fiction, in developing an account of something else, will automatically result in a fictionalist account of one's subject.[16] The second kind of problem that can result from not adequately distinguishing fictionalism from philosophy of fiction is the framing and pursuing of misguided inquiry; for example, whether some theory of literary fictions "will prove canonical for minimally adequate accounts of the other kinds of fictions – mathematical, scientific, legal, ethical, metaphysical and epistemological".[17] We find questions such as this one to be confused, both because of the obvious differences between literary works and the other subjects listed and because it employs the notion of fiction in a highly ambiguous way that seems to slide from the idea of *works* of fiction (as they occur in literature) to more of an "entity" conception of fictions – something we find problematic. Both of these kinds of problems

[14] Evans (1982). [15] Currie (1990). Cf. Woods (2007).
[16] Bueno (2009) does this with Thomasson's (1999) philosophy of fiction in developing an account he wants to call "mathematical fictionalism". This is actually the second of two theorizing strategies Bueno calls "fictionalist", but it is the one he describes as "truly fictionalist" (p. 70). As we explain in a later subsection, it is either unmotivated or problematic to classify this account as a case of fictionalism.
[17] Woods and Rosales (2010), p. 350.

can undermine the plausibility of a would-be fictionalist account and fictionalism in general.

We have noted a contrast between merely analogizing and more integrated appeals to the notion of fiction in one's philosophical theorizing, and we used it to underwrite a distinction between two different species within the genus fictionalism: comparative fictionalism and philosophical fictionalism. Although comparative fictionalism appears to be the dominant strain of the general approach, we maintain, for reasons that we will explain below, that philosophical fictionalism is actually the more useful version.

1.2 Some important aspects of fiction (or, a minor foray into philosophy of fiction)

The first thing to be clear about regarding fiction is that it is a category or status of *discourse* (as well as of other kinds of representation, though we will focus on discourse, for simplicity).[18] While this observation factors into the two species of fictionalism identified above in different ways, it ends up having the same consequence for both. Whether an account relating to some subject matter, S, is a case of comparative fictionalism or an instance of philosophical fictionalism, it will amount to meta-theorizing about *discourse* putatively about S but will not involve theorizing directly about S. For example, approaching the topic of mathematics as either a comparative fictionalist or a philosophical fictionalist involves giving a fictionalist account of mathematical discourse.

For a comparative fictionalist, fiction being a category of discourse means that what one analogizes with it and "treats as a kind of fiction" must be discourse as well. For a philosophical fictionalist, the fact that fiction is a category of discourse means that one can apply the notion of fiction only to discourse, so in order to apply the notion of fiction in the details of one's account, one must be theorizing about discourse.[19] Thus, in either case,

[18] This should not be confused with what Woods (2007), p. 1069, calls *The Literary Primacy Thesis* – the view that literary fiction is canonical and that all other types of fiction are to be understood in terms of it. Woods takes *fiction* to be a genus, with *literary fiction* as a special case of something more general. We have no objection to this, provided one is careful to recognize fiction as a genus of representation (primarily discourse). We thus allow that literary fiction need not be understood as fundamental.

[19] As a consequence, we reject even the category of "fictional entities" and therefore disagree with Thomasson (1999) that there is a philosophical issue about fictional entities that can be divided into two questions: a metaphysical question about what fictional entities would be like, if there were any, and an ontological question about whether any fictional entities exist (see pp. 3–4). The fiction-related issues are about whether fictional names refer and in what sense (if any) sentences employing them are true. Thomasson (1999) takes this up in chapter 7. Cf. Walton (1990), pp. 388–390.

fictionalism is an approach in the philosophy of language. Metaphysical or epistemological concerns often motivate fictionalist accounts and such accounts have certain metaphysical and epistemological consequences, but one should not understand either species of fictionalism primarily in metaphysical or epistemological terms.

Another aspect of fiction is that discourse (or other kinds of representations) that counts as a work of fiction is "about" – or, as we prefer to say, *portrays* – something unreal: people, places, and events – or at least features, elements of things, and so forth. We do not intend the notion of *portraying* we employ here to have ontological implications.[20] It might be considered more of a shorthand to describe the use of certain expressions, or implications regarding their use, where this does not include there *being* some entity that "is portrayed". When we say that works of fiction portray unreal things, events, or features, this is intended as a claim about how the content of the relevant discourse (or other form of representation) relates to the real world. Works of fiction do not portray the real world (at least not entirely); that is not their purpose.

The unreality of at least some of what a work of fiction portrays points to another aspect of fiction that it is crucial to recognize in order to understand fictionalism, and, in particular, philosophical fictionalism, namely, that if taken at face value,[21] (at least some of) the claims that constitute a work of fiction would not be completely "felicitous", semantically speaking. This is not to say that such sentences would automatically be false, if given a face-value reading. It is also not to say that no reading of the relevant sentences could render them semantically felicitous, even to the point of being true.[22] It is vital to understand this about sentences from a work of fiction (that is, fictive sentences) such as

(1) Sherlock Holmes sniffed sardonically. (from *A Study in Scarlet*, chapter 2)

as well as *metafictive* claims about what a work of fiction portrays, such as

[20] We think that even when 'about' is intended in an ontologically neutral sense, it still has the flavor of a success term and thus an ontologically committing twinge.

[21] We prefer to talk in terms of taking sentences at *face value*, rather than taking them literally. We think that the expressions 'literal' and 'literally' are problematically unclear and contentious. This is especially so with the adverb 'literally' when it is combined with certain adjectives, as in 'not literally true' and 'literally false'. As will emerge in later chapters, we even think that certain expression have no literal meaning (on one understanding of 'literal') because their standard linguistic functioning involves pretense. As Yablo (2000), pp. 223–224 notes, however, standard usage is easily mistaken for literal usage, further confusing the issue. The idea of a face-value reading, however, is meant to respect a combination of surface grammatical form and treatment in accepted inferential and linguistic practices.

[22] Lewis (1978[1983]), p. 263.

(2) Sherlock Holmes is a detective.

and sentences expressing their direct consequences (or presuppositions?), such as

(3) Sherlock Holmes is a human being.

for understanding the possible (perhaps hypothetical) statuses of the claims from a fragment of discourse for which one proposes a fictionalist account.

Sentences like (1) through (3) would all be semantically infelicitous, if given a face-value reading. Whether they would turn out to be false, neither true nor false (perhaps because they are non-truth-apt or maybe even meaningless), or "semantically defective" (in the sense that we will explain in Chapter 5) on such a reading depends on how we should analyze sentences employing fictional names and other fictional expressions.[23]

In contrast with John Searle, Amie Thomasson, and other so-called "fictional realists", we also think that *transfictive* sentences (sometimes called *external metafictive* sentences), such as

(4) Sherlock Holmes is a fictional character.

or

(5) Sherlock Holmes was created by Arthur Conan Doyle.

would normally count as semantically infelicitous if given a face-value reading. After all, the name they employ is a fictional name, the original use of which is in the portrayal of something unreal. We thus do not think that Doyle's writing activities brought into being some sort of abstract artifactual entity, or honed in on some preexisting, abstract entity (or some non-existing, Meinongian object). There are other ways to understand how (4) and (5) function according to which they get their intuitive evaluation as true without such additional metaphysical postulation.[24] But however we should understand (4) and (5), we should at least recognize that (1) through (3) would suffer from some kind of semantic infelicity if given a face-value reading. It might turn out that fictive discourse remains semantically infelicitous,[25] but it seems that metafictive sentences like (2) can be true. For such sentences to be semantically felicitous to this extent, their semantics must involve something more than what is assumed by a simple, face-value reading.

[23] Cf. Balaguer (1998b), p. 807, fn. 8. [24] Cf. Walton (1990), chapter 10 and Everett (2005).
[25] See our discussion of various philosophies of fiction in section 1.5.

1.3 Concerns with comparative fictionalism *qua* fictionalism

To explain the problems we see comparative fictionalism generating, it will help to set out more of the elements of this species of fictionalism in order to draw out some of its consequences. First, we should note that theorists can pursue comparative fictionalism at what we can think of as two different levels, what we call *first-level comparative fictionalism* and *second-level comparative fictionalism*.

First-level comparative fictionalism compares some subject of inquiry (or, more carefully, sentences from some discourse, D, putatively about that subject) with actual instances of fiction, correlating certain facts about them. For example, Steven Wagner compares arithmetical statements to Shakespeare's *Hamlet*, noting how the play sets standards of correctness for certain claims (making, for example, 'Cordelia is a daughter to the king' false, whereas *King Lear* makes the sentence true), in explaining his thesis of the context-dependence (relative to the adoption of an appropriate set theory) of the correctness of a set-theoretic statement about number.[26] And Rosen begins his explanation of modal fictionalism by analogizing possible-worlds sentences, such as 'There is a (non-actual) possible world at which there are blue swans', directly with metafictive sentences, such as 'There is a brilliant detective at 221b Baker Street.'[27] First-level comparisons like these can be suggestive, but if comparative fictionalism is left at this level, it is too vague, and the account does not provide enough details in its explanation of the topic of theorizing. Coming from a critical direction, sticking to first-level comparison makes this brand of fictionalism the target of certain objections that are mainly beside the point, such as a cataloging of *dis*similarities between the topic of theorizing and works of fiction.[28]

Second-level comparative fictionalism involves comparing the *account* one is offering of one's focus (that is, of the claims from some fragment of discourse, D) with some *account* of fiction – that is, with some philosophy of fiction. To the extent that the philosophy of fiction is worked out in detail, this puts meat on the bones of an instance of comparative fictionalism. As examples, consider the shift Rosen makes, from first-level comparative fictionalism to second-level comparative fictionalism, when he appeals to David Lewis's account of

[26] Wagner (1982), p. 265. Daly (2008), p. 434, fn. 30, also cites Wagner's appeal to an analogy, between the indeterminacy that Benaceraff (1965) demonstrated in set-theoretic identifications of numbers and the indeterminacy of fiction – for example, in the number of children Lady Macbeth has in Shakespeare's *Macbeth* – as a means for motivating arithmetical fictionalism.

[27] Rosen (1990), pp. 330–331.

[28] Cf. Burgess (2004). Contrast Balaguer's (2011) claim that "fictionalists are not committed to the thesis that there are no important disanalogies between mathematics and fiction."

metafictive discourse as the model for modal fictionalism,[29] and Otávio Bueno's explicit modeling of his account of mathematics on Thomasson's account of fiction.[30] While we maintain that a comparative fictionalist should pursue his theorizing at the second level and appeal to a particular philosophy of fiction as a model for (or at least as illuminating the details of) the account that he is offering, this leads to several problems or concerns about the approach.

Our first worry has to do with the attitude a comparative fictionalist takes toward the philosophy of fiction that he employs. Either the theorist is relying on the philosophy of fiction he appeals to being correct *about fiction* or he is simply using the philosophy of fiction as a shortcut to the details of his account of something else and is not actually concerned with whether this philosophy of fiction is correct about fiction. In the latter case, even if the philosophy of fiction happened to be incorrect about fiction, that would have no bearing on the account that he is offering of something else.[31] But in those circumstances, it would be odd to characterize the resulting account as an instance of *fictionalism*, since, really, nothing about it hinges on any facts about fiction.

For a case of comparative fictionalism to merit the classification *fictionalism*, the theorist offering the view must be invested in the correctness of the philosophy of fiction he employs. But this option leads to other problems. Philosophy of fiction is hardly a settled or uniform domain of inquiry. This means that a would-be fictionalist who pursues comparative fictionalism takes on the additional burden of establishing the correctness of the particular philosophy of fiction he employs in order to establish the correctness of his comparative fictionalist account of his target of theorizing. Call this the *added burden problem*. It bears noting that this problem does not arise for instances of philosophical fictionalism, since making use of a notion of fiction in giving an account of the functioning of the discourse that is the target of one's theorizing does not commit one to any particular philosophy of fiction.

Another problem for comparative fictionalism that arises from the same commitment to the correctness of a particular philosophy of fiction is what we call the *exclusion problem*. The worry is that if a comparative fictionalist has to be committed to the correctness of the particular philosophy of fiction that he employs in his theorizing, then he must reject every other rival philosophy of fiction as incorrect about fiction and must view any would-be comparative fictionalist account that employs a different philosophy of fiction as not really

[29] Rosen (1990), p. 331 (cf. fn. 9). [30] Bueno (2009), p. 70.
[31] To paraphrase Yablo (2005), p. 88 (citing Burgess and Rosen [1997]), if the all-knowing Oracle of Philosophy told us that the relevant philosophy of fiction was incorrect about fiction, it would have no bearing on whether the theorist's analogous account was correct about its subject. The theorist might need a new PR campaign or in-road for his account, but its details would be unaffected.

counting as a case of fictionalism.[32] After all, whatever the merits of some other theorist's account of her subject of inquiry, its details are not (from the first theorist's perspective) details that hold for fiction, meaning that it does not actually (at least correctly) compare its subject to fiction. This holds even if the other would-be fictionalist account is about a completely different topic. For example, if Bueno's Thomasson-on-fiction-based account of mathematics counts as mathematical fictionalism, then Rosen's Lewis-on-fiction-based account of possible-worlds-talk does not count as modal fictionalism – and vice versa. While it is plausible that, given competing fictionalist accounts of some single fragment of discourse – at most, one account can be correct – the exclusion from even *counting* as fictionalism across diverse subjects seems overly strong. But this is a feature of comparative fictionalism. Hence, comparative fictionalism suffers from the exclusion problem. In contrast, again, it bears noting that philosophical fictionalism generates no exclusion problem, for this species of fictionalism does not maintain that every philosophical fictionalist account must appeal to, or apply the notion of, fiction in the same way.

A further concern we have about comparative fictionalism arises in virtue of the criteria it employs for classifying an instance of theorizing as a case of fictionalism. The issue here has to do both with how analogizing works and with what makes for good analogizing. If, as comparative fictionalism maintains, being relevantly similar to some philosophy of fiction is sufficient for counting as a fictionalist view, then every comparative fictionalist must classify the philosophy of fiction she employs as a *fictionalist* philosophy of fiction. After all, the philosophy of fiction in question has the highest possible degree of similarity to that philosophy of fiction. Moreover, given the exclusion problem, the comparative fictionalist must view every other philosophy of fiction as not only incorrect about fiction but also as *not* a fictionalist view about fiction. But it seems highly problematic for Bueno to have to classify Thomasson's philosophy of fiction as fictionalist and Kendall Walton's philosophy of fiction[33] as *not* fictionalist. Call this the *forced classification problem*. While this is a problem for comparative fictionalism, it is not one for philosophical fictionalism. The criteria that determine whether a philosophy of fiction counts as an instance of philosophical fictionalism (and thus as a fictionalist philosophy of fiction) simply do not generate the forced classification problem. The particular way that a theorist applies the notion of fiction in his theorizing about some fragment of discourse has no bearing on whether he can recognize any account of fiction – or of anything else, for that matter – as a fictionalist account. Nor does a philosophical fictionalist's endorsement of a

[32] This also leads to a further burden problem if the theorist must justify his rejection of these extant philosophies of fiction.
[33] Walton (1990).