TRUTH AND REFERENCE IN FICTION
Stavroula Glezakos

INTRODUCTION

Fiction is often characterized by way of a contrast with truth, as, for example, in the familiar couplet “Truth is always strange/ Stranger than fiction” (Byron 1824). And yet, those who would maintain that “we will always learn more about human life and human personality from novels than from scientific psychology” (Chomsky 1988: 159) hold that some truth is best encountered via fiction. The scrupulous novelist points out that her work depicts no actual person, either living or dead; nonetheless, we use names from fiction in ways that suggest that we take these names to refer. Philosophers who investigate fiction aim to reconcile such apparently incompatible phenomena, and, in general, to account for the myriad ways that we talk, think, and feel about fiction.

FICTIONAL CHARACTERS AND FICTIONAL NAMES

Of great interest to philosophers has been the way that certain statements appear to be true in virtue of fiction. F. Scott Fitzgerald’s novel The Great Gatsby (1925) contains descriptions of protagonist Jay Gatsby’s lavish West Egg mansion; it thus seems appropriate to classify the following as true and false, respectively:

(1) Jay Gatsby lives in a lavish mansion in West Egg.
(2) Jay Gatsby lives in a polyurethane yurt near Big Sur.
The most straightforward account of truth for a sentence that contains a proper name and a predicate holds that the sentence is true just in case the referent of the name stands in the appropriate relation to the property expressed by the predicate. Thus, to maintain that (1) is true, one must either identify a referent for the name “Jay Gatsby”, or provide an alternative account of the sentence’s truth.

Many philosophers, in fact, accept that the name “Jay Gatsby” refers. They disagree, however, about what it refers to, with some maintaining that the referent of “Jay Gatsby” is not a man, others, that it is a man that is not actual, and yet others, that it is something that does not exist.

**Fictional names refer to abstract objects**

According to abstract object theorists, fictional names refer to fictional characters; on their view, critical assessment of a fictional work involves reference to such characters (see Searle 1975, van Inwagen 1977, Salmon 1998, and Thomasson 1999, 2003, 2009). Consider, for example, the following sentence, which one might utter during a discussion of the literary merits of *The Great Gatsby*:

(3) Daisy is a less well-developed character than Gatsby.

These theorists maintain that accepting (3) commits one to the truth of (4):

(4) Some fictional character exists.

This is because from (3) we can infer that there exists a fictional character that is less well-developed than Gatsby, and thus, that there exists a fictional character (van Inwagen 1977). Such characters – or “creatures of fiction” – are abstract entities,
possessing (or lacking) the sorts of properties that are at issue within critical discourse, such as being well-developed, being sympathetic, or being modelled on a particular person (van Inwagen 1977: 302). In uttering (3), one refers to two such characters, and it is their relative degree of development that determines whether one’s utterance is true or false.

Matters become more complicated when we consider sentences like (1) and (2). If the referents of fictional names are abstract objects – that is, objects that are neither physical nor causally efficacious – then descriptions within novels cannot be specifying properties of those objects. And, indeed, abstract object theorists direct us to give a revised reading to any predicative sentence involving a fictional name and predicate assigned to a character within a novel. An assertive utterance of (1), for example, should be understood as the claim that in the novel The Great Gatsby, the property of living in a lavish mansion in West Egg is ascribed to the character of Gatsby, which these theorists take to be true.

Some abstract object theorists hold that fictional characters come into existence along with the particular work of fiction in which they figure (see e.g. Searle 1975, Salmon 1998, and Thomasson 1999). On their view, creation of a work of fiction involves pretended reference to persons, places and things; by means of such pretence, an author brings into being not only her novel, but also fictional persons, places and things. Once created, a character can appear in a later work by someone other than its creator, and may be portrayed in ways that depart from the depictions offered in the
original work, so long as the new author’s intentions are appropriately directed at the character from that earlier story.

In contrast, Wolterstorff (1980) argues that the authorial task of character development is best understood as providing specifications for a sort of thing, rather than creation of a particular thing. On his account, names from fictional works refer to abstract universals, or “person-kinds,” that exist prior to, and independently of, those works. One such kind consists of the conjunction of all and only the properties that are, in The Great Gatsby, assigned to Gatsby; a sentence like (1), if used to assert that the Gatsby-kind is partly constituted by the property of living in a lavish mansion in West Egg, comes out not just true, but necessarily true. Understanding characters in this way allows us to suggest, of an actual person who happens to possess a large number of the features that constitute the Gatsby-kind, that he strikes us as an instance of that kind (although, since no actual human could satisfy all and only the descriptions of Gatsby in the novel, Gatsby-hood could never actually be instantiated).

Perhaps unsurprisingly, the proposal that fictional characters are existing abstract entities has met with some objection. One concern is that it does not seem to provide for the truth of sentences like (5) and (6):

(5) Daisy has a larger wardrobe than Nick Carraway.

(6) Either Daisy has more freckles on her left leg than on her right, or she has more freckles on her right leg than on her left, or she has exactly the same number of freckles on both legs.

Since the neither the size of Daisy’s wardrobe relative to Nick’s, nor the number of freckles on Daisy’s legs, are discussed within The Great Gatsby, the sentences are false
if given the reading *in The Great Gatsby*,... And, since abstracta do not have wardrobes or freckles, the sentences will also be false if they are read as direct predications of the fictional characters Nick Carraway and Daisy Buchanan. Thus, if we think that it is possible to make true claims about characters that go beyond the descriptions provided in the fiction, we may find ourselves unsatisfied by this account.

A deeper worry is that such fictional characters appear to be precariously balanced between inconsistency and superfluity. The fictional characters that figure in the beliefs, intentions, and emotional responses of authors and readers (see Thomasson 1999; cf. Braun 2005) do not seem to be abstracta. Take, for example, one author’s description of character development as requiring “…the very hard work of feeling a character deeply, of understanding what happiness or hurt drives that character, so that it becomes impossible to lay superficial judgment on any behavior” (Silver 2010). Similarly, readers invoke characters in descriptions of their emotional responses to fiction, as in the teenage *Gatsby* enthusiast’s powerful three-paragraph essay about the disgust, attraction, and exasperation that Daisy inspires in him. If fictional characters are abstracta, they are in no way appropriate subjects of such reactions; furthermore, being (presumably) causally inert, they can neither have, nor bring about, reactions of any kind (cf. Friend 2000: 999).

Some abstract object theorists maintain that authors do not write *about* fictional characters (see, for example, Searle 1975, van Inwagen 1977 and Salmon 1998). Nonetheless, on their view, a name that appears in a fictional work can be “imported” into reality, and used to make true assertions about a character – which, according to
Salmon, is "the very same thing that it is the name of according to the story" (Salmon 1998: 80). If, however, the author’s uses of the name in the novel do not refer to a character, it is not clear how, within that novel, properties are nonetheless ascribed to that character. (According to van Inwagen, the relation of ascription of a property to a character in a work of fiction “must be taken as primitive” (1979: 306).)

Finally, taking fictional names to refer to actually existing abstract objects appears to commit us to the falsity of sentences that we reflexively assent to, like (7):

(7) Daisy Buchanan does not exist.

To this, proponents of the view respond that (7) has different uses, and that, although it is true on some of those uses, it is not true on all of them. Someone might utter the sentence to assert something that these theorists accept – for example, that there is no person Daisy Buchanan, that Daisy Buchanan is not a real human, or that there is no object that actually possesses the properties ascribed to Daisy Buchanan in the novel. Only when (7) is used to assert something like there is no Daisy Buchanan, or the name ‘Daisy Buchanan’ does not refer do these theorists object, since, on their view Daisy Buchanan (the fictional character) does exist. Thus, abstract objects theorists maintain that they can recognize true readings for (7); all that they deny is that (7) must be given a true reading.

**Fictional names refer to non-existent objects**

Meinong (1904), however, takes (7) to be straightforwardly true. He draws a distinction between being and existence, and argues that an object need not exist in
order to be the subject of thought and talk. A thing could, for example, be a male human who walks, talks, loves, and, in general, has the features attributed to Gatsby in the novel, while lacking the property of existence; that non-existent object would be the referent of the name “Jay Gatsby,” and (7) would be true in virtue of the fact that it does not exist.

Van Inwagen (1977: 299) gives a stark statement of one objection to this proposal: it appears to “necessitate an abandonment of what are commonly called ‘the laws of logic’.” That is: how can we consistently maintain that there is something that does not exist? Moreover, if the referent of a fictional name possesses all of the properties attributed to the character within the fiction, then we should expect the referent of “Gatsby” to possess the property of existence, since, in the novel, Gatsby is portrayed as an existing walking and talking man.

Parsons’ (1980) account aims to overcome these challenges (see also Zalta (1983) and Zemach (1993) for alternative accounts of non-existent objects). He distinguishes between “nuclear” properties (like red, round, sad, human, etc. – intuitively, features by means of which an object could be identified), any set of which corresponds to an object, and “extra-nuclear” properties like existence. A fictional object, Parsons maintains, possesses only the nuclear properties attributed to it within the story. Thus, sentences (1) and (7) are true, while (2), (4), (5) and (6) are false; the truth-status of (3) is less clear. For the majority of philosophers, however, the question of whether we can arrive at a formally satisfactory account of sentences by way of non-existent objects is academic, since they remain unreconciled to the
supposed distinction between being and existence: “Is Gatsby a man? Yes. If he is
pricked (or shot), does he not bleed? Yes. Is he worth more than the whole damn
bunch put together? Probably. Just remember, though: he doesn’t exist!”

Fictional names refer to possible objects

A third account of fictional names (developed in Lewis 1978) takes “Jay Gatsby”
to refer to a man who really does exist, and who walks, talks, and loves – albeit as a
denizen of some other (concrete) possible world. Straightforward readings of (7) come
out false on this view; however, since “Jay Gatsby” doesn’t refer in the actual world, (7)
can be assessed as true if it is understood as the assertion that Jay Gatsby is not actual.

According to the possible objects theorist, at least some of the stories that are
fictional in our world – including the one presented in The Great Gatsby – are, in some
other possible worlds, “told as known fact” (Lewis 1978). In each such possible world,
the name “Jay Gatsby” refers to the individual who “plays the role” of Gatsby in that
world. An utterance of (1) – which, according to Lewis, abbreviates “In The Great
Gatsby, Jay Gatsby lives in a lavish mansion in West Egg” – is true just in case in every
world w in which the Gatsby story is told as known fact, the man who plays the role of
Gatsby in w lives in a lavish mansion in West Egg. Thus, it is not really accurate to say
that the possibilist picture provides a referent for the name “Jay Gatsby”; rather, it
provides a (different) referent in each of the possible worlds determined by the story.

One advantage of taking fictional names to refer to possible objects is that it
allows us to recognize the truth of assertions that go beyond what is explicitly
predicated of characters within a fictional work. In all worlds in which the story is told as known fact, the woman who plays the role of Daisy will have a determinate number of freckles on her legs, and so (6) is true. (5) will be true, according to Lewis, just in case among the worlds in which the *Gatsby* story is told as known fact, a world in which Daisy has a larger wardrobe than Nick differs less from our actual world than any world in which Daisy does *not* have a larger wardrobe than Nick.

Challenges to the possibilist account of fictional reference fall into two categories. In the first are general concerns about modal realism: for example, that we seem to have no means by which to verify anything about the concrete possible worlds posited by Lewis, including their existence. The second variety of objection holds that, even if we grant that such possible worlds exist, we should nonetheless deny that the referents of our fictional names are to be found there. Lewis himself recognizes (1978: 263) that he gives no account of truth for sentences like (3) and (4), and that his proposal is in that respect incomplete. Furthermore, Donnellan (1970) and Kripke (1972) argue that only an object that is causally or historically connected to a name’s introduction and subsequent use can be the referent of that name; if they are correct, then only objects in the actual world are candidate referents for names from fiction. Thomasson (2009: 16) finds the possibilist’s identity conditions for fictional characters too restrictive, since they entail that if the author had ascribed to the character “so much as one different property (however trivial), she would have written about a different possible person.” While this does seem like a dramatic consequence, the possibilist will presumably not be bothered by it, since, on his view, an author who
ascribed a different property would be a different person: any scenario involving what did not actually happen is part of some other possible world.

**FICTION AND MAKE-BELIEVE**

As noted in the previous section, some theorists who hold that fictional names refer acknowledge that, when an author is writing her novel, “there is no point in debating what sort of thing” she is writing about, since she is not writing about anything (Van Inwagen: 301). But if we accept this, and if our spirits have been dampened by challenges to the referentialist views discussed in the previous section, we might well find ourselves reconsidering the enterprise of seeking referents for fictional names.

Bertrand Russell (1919: 169) warns of the folly of any such project; on his view, “...it is of the very essence of fiction that only the thoughts, feelings, etc., [of an author] and his readers are real, and that there is not, in addition to them, an objective [character].” Russell takes those who posit referents for fictional names to have lost the “robust sense of reality” that is required for successful theorizing in this area. Some recent theorists who join Russell in denying that names from fiction refer do so with a robust sense of a particular part of reality: our games of make-believe.

These theorists (most prominently, Walton 1990; see also Currie 1993) characterize works of fiction as “props” in coordinated exercises of the imagination. Indeed, it is in virtue of playing a particular role in games of imaginative pretence, rather than any intentions on the part of an author, that a text is a work of fiction:
“[t]he notion of objects serving as props is independent of and has priority over that of the action of making objects to be used as props” (Walton 1991: 428). As with all varieties of make-believe, to consider which reports about fictional content are really true, or to externally assess the truth of what is asserted within a fiction, is to make a basic mistake. This becomes clearer when we consider two children who spend a balmy summer afternoon inhabiting the roles of Antarctic explorers adrift in iceberg-infested waters. Though the children may later tell family members lurid tales of their adventures, they would surely grow impatient with anyone who insisted on being told whether the explorers had really been growing weak from hunger, and whether they had really been able to eat their fill of plankton. On the one hand: “Yes, that’s exactly what we were pretending!” And on the other: “We were just pretending – there aren’t any such explorers!”

Walton characterization of fictional ‘truth’ instead involves “there being a prescription or mandate in some context to imagine something” (1990: 39). When we follow the mandate, and imagine that what some sentence expresses is true, and that the names within it refer, we are not, however, imagining of something that it is a certain way. This is pretence, after all, and our pretendings take the form there is a man named ‘Jay Gatsby,’ and he lives in a lavish mansion in a place called ‘West Egg.’ Neither The Great Gatsby story, nor the imaginative uses to which we put it, make it the case that (1) is true, although (1), unlike (2), is fictionally true, or true in the story – just as, in the story, the name “Jay Gatsby” is a name of a man, while “Kengo Matsusaka” is not. Pretence theorists also deny that sentences like (3) and (4) are true,
or that assertive utterances of them commit us to the existence of (real) fictional characters. Instead, they maintain that those who assert or assent to utterances of these sentences are engaged in a particular variety of make-believe, in which what is pretended is that there are fictional characters (see e.g. Walton 1990, Brock 2002).

Walton’s (1990) view is that, like fictional truth and ‘real’ truth, cognition and emotion within games of pretence should be distinguished from their non-fictional counterparts. Given that I ‘believe’ that Gatsby has just been shot, it is not surprising that I ‘feel’ sad and bitter. However, it is only within the pretence that I believe that Gatsby has been shot, and that I feel sad and bitter (that is: I am not experiencing real sadness and bitterness). What is crucial to notice, according to Walton, is that the ‘feelings’ (or quasi-emotions) that are elicited by fiction are not motivating: a moviegoer may, during a film about deadly green slime, tremble and clutch his chair with sweaty hands, but he does not run out of the theater, or believe that he needs to quickly move to higher ground, as he would if he were really afraid. In contrast, Lamarque (1981) argues that someone can be genuinely frightened by the thought of green slime, even if he is not frightened that e.g. the slime will soon engulf the theater.

Walton (1990) maintains that we are sometimes directed by fiction to imagine something that we actually believe, or know, to be true. According to New (1996), this position reveals problematic elements of Walton’s portrayal of imagination and fiction. New maintains that someone who believes that New York is the most exciting city in the world, and encounters the sentence “New York is the most exciting city in the world” in a work of fiction will not imagine, or make-believe, that the sentence is true,
or that the name “New York” refers. Rather, the reader’s thought will involve the actual city and its actual place at the pinnacle of exciting cities; thus, engagement with a work of fiction cannot, according to New, be understood as an exclusively imaginative exercise. (For a richly detailed discussion of mental pretence, see Camp (2009); for discussion of empirical results linking imagination and belief, see Nichols (2004).)

The pretence account also allows for the questionable possibility that a scrupulously accurate report of events that really did occur could qualify as a piece of fiction if it came to be used as a prop for certain imaginative exercises. The reverse is true as well: a work produced as fiction would cease to be such if it no longer figured in our practices of make-believe. Currie (1990) maintains that even if those who authored the Bible did so with fictional intentions – that is, even if the Bible was authored by those who intended it to serve as a prop in games of make-believe – the fact that “sufficiently many” people do not now read biblical stories as fiction means that the Bible could not, now, be properly classified as fiction. Thus, the pretence account seems to have the counter-intuitive consequence that being a work of fiction amounts to being treated as fiction. But this cannot be correct: we can, after all, ask about a text that is identified as a piece of fiction whether it in fact is, and we have a general idea of what sorts of investigations we might carry out in order to arrive at an answer.

TRUTH IN FICTION

Our focus so far has mainly been on accounts of truth for individual sentences within, or about, fiction. But philosophers have also considered whether a fictional
work, in its entirety, could express or present a truth, and have attended to the worry that it is not possible to learn, or come to know, such truths via fiction, since it would seem that even a work that expressed a truth could not provide evidence in support of it.

Critical assessments of a work of fiction often aim to make explicit the propositions that it expresses, or that it is committed to. One such analysis of *The Great Gatsby*, for example, takes the novel to provide “profound corrective insights into the nature of American experience,” most notably, that “the American dream, stretched between a golden past and a golden future, is always betrayed by a desolate present” (Bewley 1954). Since no such sentence about the American dream appears in the novel, we need some account of how a text might present content that is not expressed by any particular sentence(s) within it.

On one view, some fictional works express philosophical theses concerning “man and his world,” which are revealed to us via examinations of “...plot, characters, dialogue, authorial interpolation, tone and themes.” (Weitz 1975, p. 31). Exactly how, and whether, the plot and additional elements might be related to such theses is a matter of much dispute among literary critics. To the extent, however, that we offer arguments in support of the correctness of our particular interpretation of a novel’s ‘deeper meaning’, or feel unease if we begin to think that what is ‘said’ by a favorite work of fiction might be objectionable, it is plausible to suggest that we do take a work to have, or be committed to, this sort of thesis (see Harcourt 2010).
Other theorists maintain that there is no content that fiction presents assertively, or ‘as true’. Beardsley (1981) for example, holds that a work of fiction might well have a meaning discernable by an interpreter, but he denies that the work could advocate for (let alone establish) that meaning. In support of this position, Lamarque and Olsen (1994) emphasize that works of fiction rarely contain anything that could be construed as arguments for theses that they are purported to advocate.

At least some of candidate truths from fiction are not, however, of the sort that we would expect to learn via arguments. Nagel (1974) and Jackson (1982) maintain that not all truths are propositional, or objectively accessible; some are essentially experiential, or subjective. These truths concern what it is like – for example, to be in a certain situation, to be part of a certain cultural group, or to be a certain kind of person – and, according to some philosophers, we can encounter such truths via fiction, and perhaps even receive confirmation of them.

Nussbaum (1985: 516) argues that we can develop our “moral perception” by actively engaging with fiction: “certain novels are, irreplaceably, works of moral philosophy.” Reading such works with sufficient appreciation is a means of becoming sufficiently sensitive, and appropriately responsive, to particular complexities that might be present in human lives. Echoing Nussbaum, those who advocate for fiction’s place within the curriculum for the helping professions argue that such practitioners must be capable of taking the perspective of those that they serve: to successfully counsel or advocate for the bereaved, for example, “…the feelings associated with bereavement…. must either be experienced first-hand, or else vicariously through writing or other media.
which deliberately attempt to invoke them, that is, which resonate with our own feelings” (Rolfe 2002: 91).

One might worry, however, about how we are to move from acknowledgment that fiction is a means of adopting a perspective other than our own, to any claim that such perspectives from fiction provide truth, or do so in ways that allow us to recognize it as such. One source of difficulty here is that a work’s effectiveness in moving the reader to take a particular perspective, and ‘see’ things from that perspective, seems in principle independent from the truth of what is viewed from that perspective.

As an example, consider Tolstoy’s powerful short novel The Kreutzer Sonata (1889), in which the central character, Pozdnyshev, explains how and why he murdered his wife and the man that he suspected of being her lover. As the work unfolds, we (along with Pozdnyshev) become ‘enlightened’ regarding the ultimate causes of his actions; he comes to understand that the love and hate that he felt for his wife “…were different aspects of one and the same animal feeling,” and he takes his own experience as a basis for concluding that “the sex passion, in whatever form it is presented, is an evil, a dreadful evil, which is to be combated and not encouraged” (p. 329). (In an Epilogue (1890) offered in response to criticism following The Kreutzer Sonata’s publication, Tolstoy echoes the ‘insights’ that his novel provides: that “…union with the object of love…never makes the attainment of the aim which is worthy of man any easier, but always impedes it,” that “carnal love” is “an animal condition, degrading to man” and that the ideal to which all should aspire is chastity.)
Surely, for many readers, the perspective on offer in *The Kreutzer Sonata* is one that leaves them cold. Philosophers (e.g. Walton 1990, Gendler 2000, Weatherson 2004) have provided varied accounts of the phenomenon of imaginative resistance – that is, readerly refusal, or inability, to follow the imaginative prescriptions within a fiction. But not all readers will resist the invitation to take the perspective of Pozdnyshev, and to see things as he does. A skilled storyteller, writing with the fevered passion of true conviction, can make it seem that the world that he is giving us is, in fact, *the* world. And so, it appears, the perspectives to which we are provided access via fiction must always be subject to extra-fictional scrutiny; we cannot do without arguments, proceeding at least in part by way of objectively accessible propositions, if we are to assess purported truth in fiction.
BIBLIOGRAPHY


