

# FICTIONALISM AS A PHASE (TO BE GROWN OUT OF)

## ABSTRACT

It will be argued that a fictionalist account, in one central sense of that ambiguous phrase, of an area of discourse is not something that can be called correct or incorrect once and for all, but rather that areas of discourse often evolve from phases for which an error theory is most appropriate, to phases for which a fictionalist account is appropriate, and on to phases for which a straightforward account is appropriate. The phenomenon will be illustrated by examples from several areas of discourse, but especially mathematics, where plausible and implausible versions of fictionalism will be distinguished.

Unlike chemists or geologists, philosophers have no official international organization to fix the meanings of their terms of art, and when a label becomes as fashionable as *fictionalism* has done, it will inevitably be found used in several conflicting senses. Anyone discussing fictionalism would therefore do well to follow the example of Eklund (2015) and begin by marking distinctions. I will do so shortly here; but first let me note the one thing that those who profess to adopt one or another kind of fictionalist view about one or another area of discourse generally have in common: They do *not* view works in the area in question in the way we view paradigmatic works of fiction.

## THE VARIETIES OF FICTIONALISM

For the most basic feature of the distinction between *fiction* and *non-fiction*, as understood by non-philosophers such as publishers and librarians, is that works of fiction are generally not drawn on as potential sources of information in making practical

decisions: *The Hound of the Baskervilles* may be a good read, but no one needing the services of a good detective wastes any time on a trip to Baker Street. Some non-fiction works are more reliable than others, as some are more relevant than others in any particular case, but the question of reliability or relevance simply does not arise for fiction, except perhaps to a limited extent in the special case of works advertised as *based on a true story*, which is to say, as partly non-fictional.

In public libraries, works of fiction are shelved in a different way, off by themselves, alphabetically by author, from non-fiction works patrons might wish to consult for practical purposes. The latter are arranged by catalogue numbers assigned by subject matter, in order to make works relevant to a given topic easy to locate, all together in one place. Thus a gardener needing to calculate the area of an irregular plot, say in order to determine how much fertilizer will be needed, can turn the 500-599 or science section, and its 510-519 mathematics subsection, and its 516 geometry subsubsection, to find textbooks and manuals containing pertinent mensuration formulas.

Now while there are many philosophers who claim to be fictionalists about mathematics, for instance, there are none who advocate that geometry books should not be consulted in this way, and that to eliminate the temptation to so consult them they should be taken off the shelves where they can now be conveniently located, and distributed among the novels, where they will be separated from each other, and harder to find. In short, no philosopher who professes to regard mathematics as fiction really does so, not in any ordinary sense of *fiction*; and the same seems to be true about professed

fictionalists about other areas of discourse. As to what extraordinary sense of the word the philosopher who applies the term *fiction* may have in mind, one must, as I said at the outset, make distinctions.

Eklund starts by distinguishing *linguistic* fictionalism, according to which utterances in a given area of discourse are best seen as attempts to do something other than express truths, from *ontological* fictionalism, according to which the objects that the area of discourse is apparently about have a status like that of fictional characters, whatever status that may be. Here I will always have in mind the linguistic thesis. Eklund next distinguishes *hermeneutic* fictionalism, a descriptive thesis about the nature or status the discourse actually does have, from *revolutionary* fictionalism, a prescriptive proposal about the nature or status the discourse optimally should be given. Here I will always have in mind the hermeneutic thesis.

This much said, I can now give a rough statement of the hypothesis about (linguistic, hermeneutic) fictionalism that I will be suggesting here: that generally when a fictionalist account is appropriate for an area of discourse, it becomes and remains so only for a transitional period that historically speaking is comparatively brief, and that accordingly one should be suspicious of claims to the effect that a fictionalist account of this, that, or the other area is more than temporarily or provisionally appropriate.

Eklund cites another distinction, between *force* and *content* fictionalism, two different kinds of theses about what participants in whatever area of discourse in questions are up to. I understand the distinction as follows. Force fictionalism holds that

the participants mean their words literally enough, but their belief in what they say is only feigned and not sincere. Content fictionalism holds that the participants mean their words only figuratively and not literally, though what they do mean they believe sincerely enough. My discussion is intended to apply to both kinds of fictionalist view, though I will sometimes word my remarks as if specifically addressed to one kind rather than the other, leaving it to the reader to work out how what I say should be reworded to fit the opposite case.

Let me begin by stating what will probably be my most contentious background assumption, as it would apply to force fictionalism. When someone says something to us, the null hypothesis is that they believe what they are saying, though needless to say this is a *defeasible* presumption, and one all too often in the end defeated, since we all too often encounter liars and bullshitters (in Harry Frankfurt's sense). My assumption is that one is not going beyond the null hypothesis as stated if instead one expresses it as the hypothesis that speakers *sincerely* believe what they are saying. For my assumption is that in order to believe something sincerely one does not have to do anything extra beyond believing it, the reason being that what *sincere* expresses is not the presence of something extra of a positive nature, but the absence of anything extra of a negative nature. Authorities reject the folk etymology according to which the first syllable of *sincere* comes from the negative Latin particle *sine* or *without*, but the mere fact that such etymological myths have gained currency testifies to a recognition on some level that *sincere* is a negative, not a positive term.

My assumption is that it is *feigned* rather than sincere belief that requires something extra: an extra expenditure of mental effort, in order to maintain a mental reservation. My assumption is further that a speaker is always at least subliminally aware of whatever reservation is in question, so that it cannot be that speakers are feigning belief while remaining totally unaware that that is what they are doing. When we engage in story-telling, for instance, we never quite forget that important fact about our performance, and if someone enters in the middle of the tale and asks with an air of surprise, “Is that really true?” we will be ready with the answer “No, of course not.”

Eklund seems to indicate that some professed fictionalists deny this, and hold that the participants in an area of discourse may be maintaining a fictionalist attitude towards it while remaining completely unaware of doing so, and even completely incapable of coming to such awareness. I can understand what might motivate taking such a line: a desire to avoid empirical disconfirmation by the powerful *phenomenological* and *autism* and other objects put forwards in (Stanley, 2001) and elsewhere. What I cannot understand is the line itself. The kind of position Eklund seems to attribute to some professed fictionalists is one that I find simply unintelligible, and my background assumption excludes it. I will always assume that (linguistic, hermeneutic) fictionalism is about speakers’ intentions and hearers’ understandings in a psychologically real sense.

Similar remarks would apply to literalness in place of sincerity, and content in place of force fictionalism. When we speak metaphorically, we remain at least subliminally aware of the literal sense behind the metaphor. If a metaphor becomes so

overused that people forget the literal meaning, as when commentators on economic matters use the grating phrase *the height of the depression*, then one is dealing with a *dead* metaphor, which is no metaphor at all, but a new literal sense sprouted by a polysemous term, perhaps destined to become eventually the *primary* sense, as older senses wither and become obsolescent and then obsolete.

### **THREE TYPES OF VIEWS: TYPE E, TYPE F, TYPE G**

Suppose it is agreed that a fictionalist account is *not* appropriate to a certain area of discourse, and that participants in it are aiming at truth. In the general case, they will exhibit a mix of success and failure in achieving their aims, and the discourse will present us with a mix of truths and falsehoods. In an exceptional case, however, some philosophical observers may wish to suggest that for some deep reason, perhaps the failure of some basic presupposition, all the participants on either side of the various debates within the area are across the board coming out with utterances that, though sincerely intended as expressions of literal truths, are as a matter of actual fact one and all untrue, and perhaps not even false, but without truth-value.

I will use the label *type E* for views that take a given area to be an exceptional case of the kind just described. Here *E* may be thought of indifferently as abbreviating either *exceptional* or the more usual expression *error theory*. I emphasize that I am speaking of a type of descriptive thesis, which may be cited by different theorists to motivate different prescriptive proposals: to eliminate the discourse in question, or to retain it with an understanding revised in one way or another.

I will henceforth use the label *type F* for views of the kind that I have heretofore been calling (linguistic, hermeneutic) *fictionalist*. And I will use the label *type G*, abbreviating *general* or *generic*, for the unremarkable kind of view that I began this section by contrasting with type E. I may tend to speak as if we had three discrete, distinct types here, but for several reasons this is a simplification or idealization, the first reason being that F has already been acknowledged to involve two subtypes, force and content fictionalism, and Eklund distinguishes yet others.

Further, there is not a stark binary contrast between believing with full sincerity what one says, and regarding it all as merely a kind make-believe. For instance, regarding a stretch of discourse as a simplification or idealization, as I have just said my own discussion so far should be regarded, is an intermediate case. At one extreme might lie my performance (a) in a presentation before a small seminar of expert specialists, trying to state my views as exactly as possible, with all the subtle refinements for which my audience has the time and patience. At the opposite extreme might lie my performance (b) in a comedy sketch of the Monty Python type, dressing up in the vestments of the Bishop of Cloyne, and between gulps of tar-water preaching his incredible immaterialist doctrine, or some parody of it, as if I were able to believe it and did so. In-between would come performances of various kinds before audiences of students, such as (a') giving a rather rough statement of a view while reserving refinements I know are required for the question-and-answer period or separate small discussion-group meetings or later stages in the students' studies, or (b') beginning a sentence with the words "Berkeley might say..."

and then slipping into the character, speaking for some time for Berkeley rather than myself, though without donning eighteenth-century clerical garb or affecting an Irish brogue. Here, if (a) is black and (b) is white, (a') and (b') are shades of gray, the former very dark, the latter very light, and there are other shades of that color around, too.

Finally, a type E account may be correct for a substantial majority, but a type F account for a significant minority. Sympathizers with moral relativism, for instance, might hold an error theory about the moral discourse of most, especially in very religious countries, but hold a fictionalist theory about the moral discourse of many, especially in more secular countries, themselves included. Extending this last point, a change may occur over time in an area of discourse, from an period whose usage is best described by a type E account, to a later period whose usage fits better with a type F account. All that is needed for this to happen is that suspicion that some underlying presupposition of the discourse fails should gradually grow from being a minority view to becoming the majority view. The stock example is provided by astronomical talk, for instance of *sunrise* and *sunset*.

Originally, talk of the sun rising was meant entirely literally. It was understood to be a matter of the sun genuinely moving upwards to a higher altitude above the plane of a flat earth. Already with Ptolemaic astronomy this view is rejected as an error. On a Ptolemaic view, the sun is indeed rightly said to *move* as it revolves around a stationary earth; but since the sun maintains a more or less constant distance from the earth's surface, and the earth is spherical, the sun maintains a more or less constant distance from

the earth's center, so that it cannot without qualification be said to move *upwards*. For on a Ptolemaic picture *up* does not denote a unique direction in the cosmos as it does for the flat-earthier, for whom *up* is the opposite of *down*, and *down* is where the turtles are.

Rather, at any given point on the earth's surface, *up* denotes the direction away from the earth's center, and this is a different direction at different surface points, so that when the sun appears to be moving upwards from one perspective, it may appear to be moving downward from points far enough east, and when the sun appears to be moving downwards from one perspective, it may appear to be moving upward from points far enough west. Objectively or absolutely, there is no up or down to its motion, and so no literal rising or sinking.

On a Copernican picture, of course, the earth is rotating and the sun is not really moving at all; but the change from geocentrism to heliocentrism is a less radical one, as regards the question of the literal truth of statements about rising or moving upward, than the change from the flat-earth to the spherical-earth picture. Ptolemaic and Copernican views were both initially held mainly among the educated, who were a minority among the literate, who were a minority among the populace. The process of heliocentrism's displacing geocentrism was probably essentially complete by the time the spread of literacy and education made any such sophisticated view widespread among the people at large. As that gradually happened, the type F interpretation of talk of sunrise and sunset spread from the learned few to the general populace, and thereby became the correct interpretation of a discourse for which a type E interpretation would formerly have been

appropriate, when the masses were still sunk in error.

This case illustrates the occasional difficulty in applying the force/content distinction. Consider the Copernicans' way of speaking to flat-earthers around them. One may be tempted to describe them as meaning by *rise* not literally *moving upwards like a balloon*, but metaphorically *appearing* as if *moving upwards like a balloon*.

Alternatively, one may be tempted to describe them as posing as fellow flat-earthers and playing at believing that things are as they appear, and hence feigning belief in a literal upward motion. (Speaking where he might be in danger of being overheard by agents of the Inquisition, Copernicus himself would presumably be playing at being a Ptolemaicist speaking to a flat-earthier, a more complicated situation still.) The distinction that exists in principle here is perhaps not all that important in practice.

My main point in connection with this example will be that fictionalism did not remain indefinitely the correct account of astronomical language. For today a type G account must be recognized to be the correct one, owing to the conventional linguistic meanings of *sunrise* and *sunset* having changed from what they were (or what the meanings of the corresponding terms in ancient languages were) in the days when the Egyptians spoke of the sun returning from the western to the eastern edge of the world by night in a boat on subterranean waters. A look at the various free dictionaries available on line shows considerable variation in their definitions of the relevant terms, but the OED, considered by many the most authoritative, gives a definition of *sunrise* that takes a disjunctive form and reads as follows:

The rising, or apparent ascent above the horizon, of the sun at the beginning of the day; the time when the sun rises, the opening of day. Also, the display of light or colour in the sky at this time.

This is as much as to say: *the real or apparent ascent of the sun*, and so on. And the second disjunct here plainly would not have been part of a definition in the times of the Pharaohs. Its insertion is the result of a change in usage from one that ascribed a real upward motion to the sun to one that more guardedly ascribes only an apparent one.

When the term *semantics* was first coined (in Bréal 1883) what it denoted was the historical study of changes in the meanings of words, attempting to frame general laws and not just collect specific etymologies. That is not how we understand *semantics* today, but this fact itself provides another instance of the tritest truism of semantics in the original sense: meanings are always shifting, as already mentioned in connection with the topic of the mortality of metaphors. Through such processes of linguistic change as metaphors dying we may move from a situation where certain formulations were not sincerely believed when taken in the literal sense they had at the time, to a situation where they are sincerely believed when taken in the literal sense they have now, and this not because there has been any substantive change in belief, but only because what counts as the literal meaning of some key term has changed. Thus a period when a type F description would have been correct may occur as a mere transitional phase between a period when a type E description would have been correct and a period when type G

description has become correct.

## PSYCHOLOGICAL AND MATHEMATICAL EXAMPLES

Traces of changes from an E-ist to a G-ist era, with presumably an F-ist transition somewhere in-between, are especially common in medical, psychological, and psychiatric terminology. As an instance, for the term *choleric* the OED gives the following definition:

Originally: a person having a predominance of cholera among the bodily humors (now *historical*). Later also: an irascible person.

It might be more revealing to say that the original, etymological meaning, was *exhibiting the disposition characteristic of an excess of yellow bile*, and the current meaning *exhibiting a hot-tempered disposition*, the connection between the two being that the once-influential, now long-abandoned theory of the four humors took hot-temperedness to be the disposition characteristic of an excess of yellow bile. The word *bilious*, by the way, underwent the same change, being at all stages a synonym for *choleric*.

The term *choleric* derives etymologically from *cholera*, meaning *yellow bile*. But as has been noted in connection with *the height of the depression*, etymologies are easily forgotten. With the original meaning, calling a hot-tempered person *choleric* was offering a mistaken diagnosis based on correctly perceived symptoms, and a type E account of the usage would have been appropriate. With the current meaning, calling a hot-tempered person *choleric* is merely calling that person *hot-tempered* in a fancier way, noting the

symptoms but hazarding no guess at etiology, and a type G account fits best. There must, one supposes, have been an intermediate period when a type F account was appropriate, when the etymology, which must be unknown to 99% of English speakers today, was not yet forgotten and still formed a component of the word's meaning, but the four-humors theory had fallen or begun to fall into disrepute. One might then still have use the word *choleric* for its value as labeling a symptom, while silently but deliberately disowning the etiological connotations still clinging to the term.

A similar shift of meaning can be seen with *choleric*'s companion terms *sanguine* and *melancholy* and *phlegmatic*, and with another set of terms for personality types, based on exploded astrological theories: *mercurial*, *martial*, *jovial*, *saturnine*. Medicine is full of terms with similar histories, from *hysteria* to *malaria*. All, I would suggest, are relics or traces or fossils of a three-step evolution E-F-G.

In a recent work (2015) I described, without the *fictionalist* terminology of the present note, a major case of a four-stage variation on the E-F-G scheme in the history of modern mathematics, with an intermediate I between F and G. The example under discussion was the use of infinitistic methods, arguments involving infinitesimals and manipulation of infinite series as if they were finite polynomials. Here, schematically speaking, there was an initial E-like period of naive, uncritical employment of such methods, with only a few outsiders taking to heart what a lapse from the ideal of mathematical rigor was involved. The most famous of these outsiders was Berkeley, who advocated an explicitly error-theoretic view, and more specifically a compensation-of-

errors view, saying to his infidel mathematician: ...*by virtue of a twofold mistake you arrive, though not at Science, yet at Truth.*

There followed an F-like period when the methods were used with fingers crossed, or with an uneasy conscience, and with a background awareness that though they usually lead to correct results, the methods in question are capable of leading to nonsense, as in the zany proof that  $0 = \frac{1}{2}$  by manipulating the alternating series  $1 - 1 + 1 - 1 + \dots$  in a naive way. This was followed by an I-period of rejection, at least in the purest of pure mathematics, of the old ways, and their replacement by more rigorous so-called  $\epsilon$ - $\delta$ -*methods*. But once the new methods became familiar, and it became a tedious but routine exercise to recast old arguments in the rigorized new form, there was a certain G-like tendency to allow the old language back in, now understanding it as a kind of convenient shorthand.

The philosophical label *fictionalism* is vague enough that some might wish to apply it to this fourth phase as well as to the second. That is a large part of the reason why I introduced my alternative label *F-ism*, which I intend to be most applicable to the period of inchoate reservations, rather than the period when everybody knows that the formulations one uses are, like current uses of *sunrise* or *choleric*, not to be understood in their etymological senses, so to speak. If one does apply the same label *fictionalist* to both periods, then one must recognize another distinction between two kinds of fictionalism to go with linguistic *versus* ontological, hermeneutic *versus* revolutionary, force *versus* content, and others such as extrinsic *versus* intrinsic to be found in the

literature. The distinctive mark of the fourth period is its *self-consciousness* and *documentability*. The official explanations of what is really supposed to be meant are something of which the participants in the discourse are fully aware, whose details can be found written down in standard textbooks, study of which is part of professional training in the relevant field.

One mathematical case of the phenomenon was striking enough that it may be worth recalling separately here, that of Poincaré's proof of his celebrated recurrence theorem (whose technical content is fortunately not material in the present context). Poincaré gave an argument in terms of probability that no one at the time, and least of all Poincaré himself, would have known how to make rigorous. After probability theory was incorporated into rigorous mathematical analysis by Kolmogorov and others, however, it became possible to read the words of argument against a background of tacitly understood rigorous definitions that Poincaré himself had been in no position to supply, and find in it a perfectly rigorous proof. I compared the case to the Borges story about an author who in the twentieth century rewrote (parts at least of) *Don Quixote*. Borges quotes passages from the famous Cervantes and his fictitious Menard that are word-for-word the same, commenting on how much deeper is the version written after, rather than before, William James.

Physics provides an unlimited number of further examples of four-stage developments: (E) naive acceptance of a theory that we retrospectively recognize is not entirely correct, (F) continued use of the theory despite doubts and reservations

occasioned by known anomalies, (I) development of a new theory officially superseding the old, (G) continued use of the old theory where the new one predicts that it will produce good approximations, in the documentable understanding that what one really means is not that the old-style formulations are exactly correct, but that they are acceptable approximations. The relationship between Newtonian and relativistic gravitational theory is one of many examples. In some areas we are still stuck at stage F; renormalization may be an example. But everyone hopes and almost everyone believes that the present stage will be temporary.

It is in the nature of science that its practitioners should acknowledge their fallibility, and search out possible errors, and that when faults are found in an accepted theory, they should be made known publicly (or at least privately if political and religious authorities make public acknowledgment too dangerous). Thus it is in the nature of science that E-periods should be followed by F-periods.

It is also in the nature of science that when reservations arise about an existing theory, the search for a new theory that would correct the errors of the old, explain its past successes despite the presence of error, and permit calculation of the scope and limits of its safe future applicability, should get put down on the science's to-do list. It may not be feasible at a given moment to do much more than note the anomalies, but they will have to be dealt with sooner or later. Thus it is in the nature of science that F-like periods should be succeeded by G-like periods in the manner already described for astronomy, medicine, mathematics, and so on.

Some philosophers may wish to apply the *fictionalist* label to both periods, but if so, the fictionalism of the G-like period will be distinguished by being something not only conscious but documentable. My suggestion in view of the various distinctions and examples that have been put forward is that we should be highly suspicious of fictionalist accounts that would make the fictionalistic character of a scientific discourse a state of affairs that is *permanent* yet not *documentable*. And as a life-long anti-nominalist I will predictably cite most current versions of mathematical fictionalism as views that ought to be regarded with the strongest suspicion on just such grounds.

The point is too large to debate here, and in any case the anti-nominalist diagnosis of the nominalist's error has been available since (Carnap 1950). Mathematicians have more or less definite standards for evaluating purported proofs of existence theorems. By these standards, some existence questions, for instance, the problem whether there are any odd perfect numbers, are difficult, while others are easy. The problem whether there are any even prime numbers greater than two has an easy negative solution, and the problem whether there are any prime numbers at all greater than two has an easy positive solution. Nominalists reject established scientific standards of evaluation, and being left without any agreed alternative standards, they are left with interminable debate over what Carnap would call a futile pseudo-problem.

The distinctive error of the *fictionalist* nominalist is to try claim support from the attitudes of mathematicians, and somehow parlay the fact that algebraists and geometers generally have no knowledge of or interest in the supposed extra-mathematical,

philosophical sense of existence debated by ontological metaphysicians, into a point for the nominalist side in that debate. Often this attempt is accompanied by an error about the meaning of *literal* parallel to the error about the meaning of *sincere* that I warned against earlier: the error of supposing it connotes the presence of something positive rather than the absence of anything negative. As a result, the failure to detect anything extra one way or the other is misinterpreted as evidence that mathematicians mean what they say only in a figurative and not a literal sense, or are agnostic between the two. The chief value of historical examples of genuine but transient F-ist periods is perhaps as a foil to the spurious F-ism that results from such misinterpretations.

## REFERENCES

- Bréal, Michel (1883) “Les Lois Intellectuelles du Langage: Fragment de Sémantique,”  
*Annuaire de l’Association pour l’Encouragement des Études Grecques en France*  
17:132-142.
- Burgess, John P. (2015) *Rigor and Structure*, Oxford: Oxford University Press.
- Carnap, Rudolf (1950) “Empiricism, Semantics, and Ontology,” *Revue Internationale de Philosophie* 4:20-40.
- Eklund, Matti (2015) “Fictionalism,” in E. Zalta (ed.) *Stanford Encyclopedia of Philosophy*, winter 2015 ed., URL =  
«<http://plato.stanford.edu/archives/win2015/entries/fictionalism/>».
- Stanley, Jason (2001) “Hermeneutic Fictionalism,” in P. French and H. Wettstein (eds.),  
*Midwest Studies in Philosophy XXV: Figurative Language*, Oxford: Blackwell,  
36-71.

## BIOGRAPHICAL NOTE

John P. Burgess is John N. Woodhull Professor of Philosophy at Princeton University, where he has taught for forty years. He has published eight books and numerous articles in logic, philosophy of mathematics and language, and history of analytic philosophy.

John P. Burgess  
Department of Philosophy  
Princeton University  
Princeton, NJ 08544-1006 USA  
[jburgess@princeton.edu](mailto:jburgess@princeton.edu)