

Piotr Brykczyński

## **Aiming at Truth**

### **Part II**

#### **8. THE *AS IF*-SENTENCES AND IMITATION**

Consider the *as if*-sentences. They are composed of the *as if*-conjunction and two sentential components. The second of them forms with the *as if*-conjunction an *as if*-clause. The *as if*-sentences share important features of meaning with subjunctive conditionals, and, to a lesser degree, with sentences distinctive of the causal and dispositional talk. A feature that is shared by the meaning of sentences of all these categories is a modal commitment of a sort: the truth conditions of these sentences involve necessitation. Strictly speaking, it is difficult to definitely settle the question of whether one can speak of necessitation with reference to the features of meaning of sentences of all the above categories unambiguously. Since, however, there is no good evidence for ambiguity, I shall assume its absence tentatively, following the rule: meanings are not to be multiplied beyond necessity.

The above similarities in meaning establish a wide area of fruitful cooperation between the respective branches of semantic analysis, including research on their ontological foundations. One of the results are wide-range reductive accounts with the language of subjunctive conditionals as supplying paraphrases (for the prospects of application to dispositional talk, cf. Bird 2007: 24-42). In view of such cooperation and reductions, my account of the semantic properties of the *as if*-sentences may look somewhat isolated, but it only aims to convey some guiding ideas, which seems to make its degree of isolation tolerable.

If there are states of affairs, they have manifestations (relative to certain circumstances, e.g. to 'normal' circumstances, or any possible circumstances). Some other states of affairs manifest them (e.g. the state of affairs: that this street here is now

wet, manifests the state of affairs: that it rained here a few minutes ago). Manifesting a state of affairs relative to any possible circumstances is tantamount to being necessitated by it. My use of the term “manifestation” is closely akin to its use as a technical term of the language of the theory of dispositions. One minor point to be made in this connection is that manifestations of a disposition tend to be conceived as events rather than as states of affairs. However, the purpose of the remarks to follow makes the difference negligible. Leaving it aside, observe that according to the so called (simple) conditional analysis of predication sentences with overtly dispositional predicates (as “is disposed to break when struck”) the functioning of a disposition consists in that its manifestations are necessitated by the respective stimuli. However, what is usually diagnosed as a stimulus relative to a certain disposition does not necessitate the respective manifestation. There is perhaps no conceptual misuse in such diagnoses. It is an open question (cf. Bird 2007: 39). Arguably, the pattern of necessitation distinctive of the functioning of dispositions is much more complex than expected on the above grounds. But in noticing it we do not commit ourselves to abandoning the view that when a disposition is *at work*, its being at work consists in necessitation of its manifestations (note: when a disposition has not been put to work, necessitation is present potentially). Thus the way the term “manifestation” has been introduced above exhibits no significant disagreement with the way it is used as a technical term in dispositional talk.

If there are states of affairs, there are states of affairs that consist in a certain person being in pain at a certain time. Call them ‘states of affairs of being in pain’. If, further, there are types, there are types of states of affairs, one of them being the type of states of affairs of being in pain (note: the type of *S*-s is instantiated by *S*-s, each of them instantiating it in virtue of being an *S*). Let  $S_T$  be a variable ranging over types of states of affairs, and let  $\text{neg-}S_T$  be the type of states of affairs that is negative with respect to  $S_T$  in the sense of being the type of the respective negative states of affairs (suppose there are negative states of affairs). Consider types of facial expression such that their instantiation is determined in a purely behavioural way. There is probably — among types of facial expression that meet this condition — no type  $T$  such that the states of affairs consisting in having a facial expression of the type  $T$  are *bound* to be manifestations of the states of affairs of being in pain. But we can approximate this relationship practically without limit. And when a certain level of proximity is reached, we can say that the given type of states of affairs of having a certain facial expression is *diagnostic for* the type of states of affairs of being in pain. To turn to a general account, let  $S_T$  be a type of states of affairs. Types of states of affairs differ in respect of the degree to which their instances are bound to be manifestations of the states of affairs of the type  $S_T$ . For a type of states of affairs to be *diagnostic* of  $S_T$  it is sufficient and necessary that a certain limit value be reached (it is determined with vagueness and context-sensitivity). One can characterize the property of being perfectly diagnostic, strongly diagnostic, weakly diagnostic, etc., in a similar way.

Consider again the type of states of affairs of being in pain. Call it  $S_p$  for short. A person with a facial expression *typical* of being in pain may nevertheless *deny* being in pain, some sorts of such denial determining types of states of affairs that are diagnostic of the type  $\text{neg-}S_p$ . It shows that an action can involve both instances of a type of states of affairs diagnostic of a certain type  $S_T$ , and instances of a type of states of affairs diagnostic of  $\text{neg-}S_T$  (of course, of such two cases of being diagnostic at least one will not be a case of being perfectly diagnostic). Call such action ‘diagnostically inconsistent’ with respect to the given type of states of affairs (in the above example: the type of states of affairs of being in pain).

To return to the *as if*-sentences, suppose someone says:

- (1) He acts as if he was in pain.

What is conveyed can be characterized partly (and roughly) in terms of the relation of being diagnostic as holding between types of states of affairs: if someone acts *as if* in pain, the given action is diagnostically consistent and involves instances of a certain type of states of affairs that is diagnostic of the type of states of affairs of being in pain. We may call the respective features of meaning of the *as if*-conjunction the ‘diagnostic consistency feature’ and the ‘diagnostic sufficiency feature’. So far so good (if the level of approximation is tolerable). But now suppose someone says in the same breath:

- (2) He acts as if he was in pain. And, as a matter of fact, he is in pain.

It may look awkward to say something like this, which may suggest that the *as if*-clause in (1) is counterfactual (it would make (2) contradictory).<sup>1</sup> But it is perfectly natural to say:

- (3) He acts as if he was in pain. But he isn’t.

If the *as if*-clause in (1) was counterfactual, (1) would entail the given person is not in pain, this entailment making (3) pragmatically unaccountable. The problem we thus face is this: if the *as if*-clauses are not counterfactual, then what makes (2) awkward? Call a state of mind ‘expressed’ by a sentence, if the proper use of this sentence to make a statement (the use that does not violate the rules of language) requires that the speaker be, while using it to make a statement, in this very state. What is expressed in this sense may not affect what is stated (asserted), as in the case of some emotive features in meaning. Thus some sentences are used to make a statement with an *expressing vs stating discrepancy*.<sup>2</sup> As regards the *as if*-clauses, it may

<sup>1</sup> One could also speak of the whole *as if*-sentences as counterfactual with respect to their *as if*-clauses. The choice involves a trade-off between semantic motivation and brevity.

<sup>2</sup> For the above concept of expressing and the distinction between expressing with and without reflection in what is stated, cf. especially Ajdukiewicz 1956/1978. The question of the nature of factors responsible for the expressing vs stating discrepancy is intimately related to the main issues about presuppositions and conventional implicature. Needless to add that the conceptual clarity in

be maintained (and it will be assumed tentatively in what follows) that their proper use requires (let us neglect the dependence on tense form) that the speaker be in the state of mind of not believing that the given state of affairs obtains (in the above example: the state of affairs of the given person being in pain), this feature not being reflected in what is stated by the corresponding *as-if* sentences (observe that it would be thus reflected if the *as if*-clauses were counterfactual). It follows that the *as if*-sentences are used with the expressing *vs* stating discrepancy.<sup>3</sup>

Linguistic means responsible for the expressing *vs* stating discrepancy are somewhat defective in respect of drawing distinctions. They may be and often are disabled from being used to draw distinctions in a way accessible to all language users. Now, on the above interpretation of the *as if*-clauses, they suffer from being defective in this very way. It does not have analogy in speaking of imitating. While what is thus conveyed seems to exhibit a close analogy to what is conveyed by the *as if*-sentences, in respect of the way being diagnostic is involved, speaking of imitating involves a counterfactual component and displays no expressing *vs* stating discrepancy. Therefore we can distinguish, without the above limitations, the universe of cases of imitating, and we can distinguish between various types thereof, including distinctions that are crucial for the methodology of positing fiction (as the distinction between functional and dysfunctional cases).

According to the above account of meaning of the *as if*-sentences, if, e.g., someone acts as if in pain, this action is to some extent, in respect of involving or not involving manifestations of the states of affairs of being in pain, *as* some actual or possible actions that involve such manifestations. It does not follow that the *as if*-statements are elliptic for some *as*-statements. What is more, the use of “as” as occurring in the *as if*-conjunction can be accounted for without the hypothesis of ellipsis. It is enough if the *as if*-sentences obviously entail some *as*-sentences, which, as we have seen, is the case.

In characterizing the meaning of the *as if*-sentences as involving what has been called the ‘diagnostic sufficiency feature’, I follow to some extent Vaihinger’s analysis (1924: 91-95, 256-266). Moreover, there seems to be no reason to think that Vaihinger would not acknowledge the diagnostic consistency feature. On the other hand, he regards the *as if*-clauses as counterfactual, and he regards the *as if*-statements as elliptic for certain *as*-statements.

The above comparison with Vaihinger’s account of the *as if*-sentences neglects his examples, in contradistinction to what he has to say generally about them. But there is an interesting discrepancy between what Vaihinger has to say generally about his examples, and what they themselves say when we let them speak for themselves. One of Vaihinger’s examples is: “Matter is to be regarded *as if* it consisted of atoms”

---

relevant research leaves much to be desired. The concept of stating is not exceptional in this respect.

<sup>3</sup> For a possible analogy in the use of subjunctive conditionals or some sort thereof, cf. Bennett 2003: 11-12.

(1924: 82). This sentence can be paraphrased as: “It is obligatory (desirable, advisable?) that matter be regarded as if it consisted of atoms”. Observe that if what is at issue is the meaning of the *as if*-sentences, what counts directly as an example is the sentence that makes up the scope of the obligation operator (more exactly: its counterpart with the subjunctive of the verb “be” replaced by the indicative). Now, some states of affairs consisting in our ‘regarding matter’ in this or that way involve beliefs concerning the constitution of matter. It would be rather artificial, to say the least, to speak of all such states of affairs as *manifesting* the given beliefs. But there are perhaps exceptions. On the other hand, the states of affairs consisting in our ‘regarding matter’ in this or that way cannot be said to manifest the states of affairs the beliefs they involve are *about*. It would be a plain absurdity to say, e.g., that certain states of affairs consisting in our regarding matter in this or that way manifest the state of affairs: that matter consists of atoms.<sup>4</sup>

Thus Vaihinger’s example does not back up the above analysis. It does not back up his own analysis either. And it is by no means an exception. Most, if not all Vaihinger’s examples disagree with the above analysis and with his own analysis (which to some extent comes to the same thing) in an analogous way (cf. e.g. Vaihinger 1924: xlvii, 86). What is more, the examples one meets in the present-day works on fictionalism (some of them at least; see e.g. Elgin 2009: 78 and Fine 1993: 34-35) do not differ from those to be found in Vaihinger in respect of what is at issue here. Thus we face a serious interpretive difficulty. What can we do with it? Is the above analysis, which in the relevant respects is also Vaihinger’s analysis, incorrect? In view of how widely the above type of examples is represented, ‘the general account vs examples’ discrepancy in Vaihinger should not be accounted for in terms of misuse. We are rather dealing here with two meanings of the *as if*-conjunction, one responsible for Vaihinger’s general account, and one responsible for his examples. These two meanings, observe further, are likely to be genetically linked by a shift in meaning. But what, exactly, such shift in meaning could consist in, and what could be its rationale?

Our activities directed by having a certain belief (and thus manifesting it) need not be merely cognitive activities, and those that are, can display various distance to methodological standards, depending on our methodological competence and other circumstances. Thus an imitation of having a certain belief need not be confined to imitation that takes place in the realm of cognitive activity conforming to methodological standards. But in the methodology of positing fiction we may be interested in imitation with such restriction. It has to be noticed in this connection that the ‘epistemic descent’<sup>5</sup> from “as if we believed that *p*” to “as if it was the case that *p*” is apt to bring

---

<sup>4</sup> Of course, on some views on the mind–body problem mental states of affairs can be said to manifest certain states of affairs that determine the constitution of matter, but the mind–matter relationships responsible for these purported instances of the manifestation relation are not instances of aboutness (and establish different correlations).

<sup>5</sup> Cf. the use of the term ‘semantic ascent’ for analogy that elucidates what is meant here.

about, as a by-product of a sort, a change of focus that roughly agrees in direction with making the above restriction (the relevant direction is that of approaching objectivity). Thus it may be conjectured that in some contexts the epistemic descent in question conventionally serves the purpose of making this restriction in a succinct way. Not without some additional merits. The shift from thoughts to what they are about can be accompanied by relaxing the conditions concerning the character of the cognitive states that are considered as manifested in our cognitive activity. Thus we may leave out of account the distinction between beliefs and conjectures. Observe, further, that if one of the above two meanings of the *as if*-conjunction has been acquired by way of a shift in meaning, it is rather the one indicated later. It follows that the shift in question involves a shift from beliefs and conjectures to what they are about. Therefore we may call it the ‘objectual shift’.

The difference between the ‘normal’ use of the *as if*-conjunction and its use with the objectual shift can be made more salient by a certain comparison. Once again speaking of imitation turns out comparatively relevant. Recall the similarities and differences that have been indicated in the case of comparison of the concept of imitation and the ‘normal’ use of the *as if*-conjunction. When we turn to the objectual shift, new essential differences appear: it does not make any sense to speak of, e.g., our actions as imitating matter’s consisting of atoms.

The objectual shift is not the only shift in meaning of the *as if*-conjunction. There is another one. It consists in introducing a certain relativization. Suppose I say: “They act as if they believed that matter consists of atoms”. It would be quite natural to object, pointing to some features of the given persons’ non-cognitive actions that betray them as not believing. And it would be equally natural on my part to respond that what was meant was only imitation in respect of manifestations that belong to *cognitive* activity. The adequacy of such response requires that my statement be elliptic, the expanded sentence containing some constituents that express an appropriate relativization (the latter has to be analogous to the relativization of imitation in my response). Now, to use the *as if*-conjunction with this relativization is to use it with a change (a shift) in meaning. We may call this shift in the meaning of the *as if*-conjunction the ‘relativization shift’. The relativization shift allows to speak of imitation with restriction to imitation in this or that respect, including the restriction indicated above as illustrative.

The use of the *as if*-conjunction with the objectual shift has some advantages over speaking of imitation of states of having a certain belief or making a certain conjecture: speaking of imitation does not allow (without additional linguistic devices) equal flexibility in respect of propositional attitudes we take into account as involved in what is imitated. However, as soon as we make restriction to, say, beliefs, this advantage can be neglected. Thus with such restriction we can dispense with the use of the *as if*-sentences altogether, speaking of imitation instead. Of course, in what follows we will be concerned only with imitation of the states of having certain beliefs (‘belief imitation’ for short) in respect of manifestations that belong to cognitive activity.

## 9. BELIEF IMITATION IN SCIENTIFIC PRACTICE

### 9.1. Make-believe and other types of belief imitation

Recall that beliefs can be understood as dispositional states. Instead of speaking of beliefs understood this way, one can speak, to avoid ambiguity, of (dispositional) states of having a belief (incidentally, having a belief may also be understood as a dispositional *property*). That is how beliefs are understood in the present paper. Beliefs understood as dispositional states have to be distinguished from mental occurrences that manifest them. The latter belong to the category of judgements (all judgements are mental occurrences of a sort). The belief in the proposition that  $p$  will usually manifest in some circumstances in judgements that  $p$ . However, in extraordinary circumstances we can make judgements that contradict our beliefs.<sup>6</sup> What is more, we can temporarily live in a world of make-believe. While it is not correct to say that we then have *temporary* beliefs that contradict our beliefs *tout court*, such way of speaking is not totally unnatural either, and it can easily be adopted by stipulation.

It has to be noticed in this connection that having the temporary belief that  $p$ , while living in a world of make-believe, is a special form of imitation of having the belief that  $p$ .<sup>7</sup> This kind of belief imitation is ubiquitous in the research practice in science.<sup>8</sup> It is not necessary to be a professional philosopher, let alone a metaphysician, to deny the existence of certain posits even if they are posits of well established theories. But it would appear as a hindrance to our cognitive activity, with no well visible benefits in return, to have a constant awareness of the facts that make such denial legitimate. Anyway, more often than not we live in the ‘world’ of our theoretical fiction in the state of oblivion (it becomes then a ‘world’ of make-believe), letting some methodologically extraordinary circumstances to be reminders. Such circumstances may result, e.g., from a change of theoretical setting into one more demanding in respect of exactitude. Another possible source is a shift from a physical (or biological, etc.) to a distinctively metaphysical perspective.

Note that whatever the level of language of the theories we avail ourselves of in our research practice in some scientific discipline, such research practice involves availing ourselves also, for methodological reasons, of the corresponding metathe-

---

<sup>6</sup> More exactly: judgements expressed by sentences that are contradictory with respect to sentences expressing our beliefs.

<sup>7</sup> Not all dispositions whose manifestations are judgements belong to the category of beliefs. Beliefs have special stimulus conditions. Temporary beliefs are not beliefs. They manifest in judgements but differ from beliefs in stimulus conditions. These differences, together with the character of manifestations, make them responsible for *temporary* belief imitation (hence the term “temporary beliefs”).

<sup>8</sup> For the importance of make-believe in the research practice, cf. Barberousse, Ludwig 2009: 58, Nolt 1986, Suárez 2009a: 10-11, 14-15, and Yablo 2005: 98-100.

ories, including metatheories that take into account the semantic dimensions of the given theories. The shift to such metatheories usually preserves (to some extent at least) the troublesome ontological commitment. Thus, a shift to a metatheory does not of itself preclude remaining in a world of make-believe, independently of the methodological status of the distinctively linguistic posits.<sup>9</sup>

Obviously, temporary living in a world of make-believe in science is not a delusion or self-deception. This is not to say that psychological research on these phenomena affords no prospects for fruitful exchange of ideas. It is worth noting in this connection that some recent accounts of delusion and self-deception in psychology take issue with the idea that these phenomena involve flouting norms of belief formation (Bayne, Fernández 2009: 8, 17). It has been suggested, in particular, that “delusional patients sometimes engage in a kind of double bookkeeping, in which they confine their delusional fantasies to a world of make-believe” (Bayne, Fernández 2009: 8). Moreover, as Bayne and Fernández point out, “in many cases of self-deception the subject is aware of truth ‘deep down’, as it were” (2009: 11). According to the authors, such awareness bears on the constraints to be imposed on theories of self-deception. One of the conditions a theory of self-deception has to meet is, as they put it, “that it account for the fact that the subject appears to know the truth *at some level*” (2009: 12; italics mine). Now, it is not obviously illegitimate to say that in living in a world of make-believe *in science* one ‘appears to know the truth at some level’. But how, exactly, are we to make sense of the metaphor of levels here? One of the guiding ideas should be, I suppose, that speaking of differences of level can be replaced to some extent by speaking of differences in strength of the stimulus conditions of dispositions that share with beliefs the feature of having judgments as manifestations.

When some methodologically extraordinary circumstances act as reminders, expelling us from the worlds of make-believe we temporarily live in, we deny the existence of the objects ‘responsible’ for the expulsion. When this happens while we avail ourselves of some theory with such objects among its posits, there will usually be some continuity: the formal component of this theory will be used on a different basis. This continuation may consist to a degree in the assertion being withdrawn without changes in the objective content of thought. Another form of continuation is involved in the appropriate shifts to metatheory. One of the options is also a purely instrumental use of the formal component of the given theory (cf. section 12 for the methodological appraisal of this option). Incidentally, it may be observed that the purely instrumental use of the formal component of theories depends for its effectiveness on the extent to which they have been formalized. Now, to mention formalization in connection with what is at issue here is not to deny that there are sources of

---

<sup>9</sup> The account of acceptance (of theories) that has been put forward tentatively in section 1 is perhaps in need of correction allowing for make-believe on the metalevel. I shall not elaborate on this point.



usefulness of formalization which have nothing to do with the ontological status of what is posited. As a matter of fact, the best known source of usefulness of formalization is of this very sort.

One can gain an additional insight into the methodologically significant diversity of belief imitation by answering the question of what can betray us as non-believers. Instances of belief imitation can differ interestingly in this respect. These differences are accompanied by differences in relationships between the mental state we are in and the state of having the given beliefs. Due to the character of the differences of the second kind, the imitation can be viewed upon as more or less superficial in certain important respects. Of course, being superficial in a certain respect need not imply being superficial in all other respects. One should not expect a perfect uniformity of grading. Another point to be made here is that there are probably no methodologically interesting differences between instances of belief imitation that are not accompanied by interesting differences in what betrays us as non-believers. It goes without saying that some differences of the latter sort may themselves be worth methodological consideration.

One of the types of belief imitation that deserves special mention in methodological settings is the type of imitation distinctive of taking a crypto-metatheoretical perspective. Worthy of separate mention is also the type of belief imitation distinctive of purely instrumental use of the formal component of theories. There is some overlapping of instantiations here, which calls for joint mention. As regards other possibilities, the above grading in respect of superficiality of imitation makes it natural to mention in the first place cases of belief imitation involved in the uses of the formalism of a theory in a way resulting from a reinterpretation. Observe that in using sentences as having meanings established by reinterpretation we need not use them with assent. If we do use them in this way, the belief that is expressed is different from the one they express in 'normal' conditions.

Last but not least, we have to distinguish cases of living temporarily in a world of make-believe. Whereas cases of use with a certain meaning established by reinterpretation instantiate a certain *façon de parler*, in considering some case of imitation involved in living in a world of make-believe we are dealing rather with a certain *façon de penser*. The comparison with literary fiction seems instructive in this connection (Nolt 1986, Kalderon 2005: 3, Loux 2006, 80-81).

The foregoing discussion in this section has left in the background the distinction between fictionalist and non-fictionalist solutions to the problem of theoretical fiction. Now, what is most characteristic of fictionalism (though not necessarily common to all versions thereof; cf. below and section 7) is endorsing the belief imitation involved in living in worlds of make-believe. But fictionalism does not demand that we imprison ourselves in such worlds. Granted, fictionalism does not endorse belief imitation that is distinctive of making use of theories in a way that results from reinterpretation intended as eliminating the troublesome ontological commitment. The reason is simply that fictionalism denies the need for such reinterpretation. As re-

gards the other kinds of belief imitation that have been distinguished, there is no absolute restriction. What sorts of belief imitation can be endorsed depends on circumstances.

As has been said in the preceding paragraph, the most characteristic trait of fictionalism is endorsing belief imitation involved in living in the worlds of make-believe. It does not follow that one cannot be a fictionalist without taking this position. Fictionalism admits of versions that oppose it.<sup>10</sup> They are relatively close to instrumentalism.

## 9.2. Make-believe and games

In Nolt's paper *What Are Possible Worlds?* we are asked to treat the possible worlds discourse in ontology and semantics as "*playing a game of »make-believe«*" (1986: 435; italics mine). There are, of course, cases of make-believe characteristic of playing games. Perhaps one can also distinguish among games 'games of make-believe'. But when we read a novel we do not play a game. Nor do we play a game when we engage in make-believe in science. Granted, temporary living in worlds of make-believe in science shares some important traits with game-playing. The workings of imagination and creativity in the first place. Nevertheless, it *is not* playing a game. Nor has it stages that could be characterized in this way. It is perhaps enough to notice functional differences to make this claim obvious. To return to Nolt's paper, a literal interpretation of the passage that has been referred to would be perhaps a misinterpretation. The more so that the suggested connection with games looks somewhat looser when the view is expressed that "the *paradigm* by which possibilistic discourse is understood, should be games of make-believe" (1986: 440; italics mine). Likewise, a looser connection is suggested when the discourse Nolt is concerned with is characterized as playing a 'language-game' (1986: 440, 444). Since games are a paradigm of a sort according to Wittgenstein's conception of language-games, this second refinement in Nolt's conception of 'playing a game of make-believe' goes hand in hand with the first one.

It is worth noting in this connection that Wittgenstein's conception of language-games aims, *inter alia*, at highlighting the "multiplicity of the tools in language and of the ways they are used" (Wittgenstein 2004: §23).<sup>11</sup> Since fictionalism helps to reveal such multiplicity, Nolt's considerations can be regarded as adding an interesting new brush stroke to the fictionalist picture. On the other hand, the 'game-discourse' in philosophy ('language-games discourse' in particular) initiated by Wittgenstein's late philosophy may often bring about confusion, and the question arises whether the benefits, including those indicated above, still outweigh the costs. What was a nov-

<sup>10</sup> For the question of whether Field's fictionalism is such a version, cf. section 7.

<sup>11</sup> I am not the first to refer to this paragraph of *Philosophical Investigations* in connection with the way fictionalism reveals functional multiplicity in the use of language; cf. Kalderon 2005: 5.

elty in the thirties and forties of the 20<sup>th</sup> century need not be a novelty now. This is not to say that making comparisons with games and using the metaphor of game exhausted their potential to afford new insights, but the theoretical settings that can make it manifested change. As regards make-believe, some authors tend to use the terms “make-believe” and “game of make-believe” almost interchangeably.<sup>12</sup> I am not inclined to follow this practice.

### 9.3. Content-oriented and prop-oriented make-believe

Toys, paintings, and novels stimulate our imaginations leading us into worlds of make-believe (Walton 1993: 65). Moreover, they impose restrictions on how the given worlds of make-believe are shaped (Yablo 2005: 96-96) and “guide our travels through them” (Walton 1993: 65). Such aids in living in a world of make-believe are called ‘props’ (Walton 1993: 65, Yablo 2005: 96).<sup>13</sup>

Theories can function as props. Moreover, when some of their posits are regarded as real and shape the given world of make-believe ‘from outside’ (with respect to what is imposed by the given theory alone), they are props too.<sup>14</sup> However, we should rather not expect that our answers to the question of how reality and fiction mix in the worlds of make-believe that are shaped by our theories will reveal something very instructive. They are not likely to show something decidedly new. Speaking of props seems to be more fruitful e.g. in the case of illustrations in scientific books or arrow diagrams in category theory (for a general appraisal of the role of arrow diagrams, cf. Corry 2004: 362-363). Of course, they are not props unless the worlds they lead us into are worlds of make-believe.

When living in a world of make-believe is an end for itself (as in games and in reading novels), a prop acts only as a means to an end. But for an object to be a prop that leads us into a world of make-believe, it has to stand to the mental representation of this ‘world’ in certain relationships which make possible a change in roles that is almost a reversal: we can let ourselves to live in the given world of make-believe to gain information about the respective props. So it happens, e.g., when we let ourselves being ‘caught up’ in a story in order to achieve better understanding of the content of the given literary work. With reference to Walton (1993: 67), Yablo (2005: 98) makes this point generally in the following way:

<sup>12</sup> Cf. Walton 2005 and Yablo 2005; perhaps we should also mention Nolt 1986 in this connection.

<sup>13</sup> What is intended is explained in a very rough way. Moreover, while Walton brings to the foreground the role of stimulating imagination, in Yablo one can get at first sight the impression that it is only the other of the two functions that is directly relevant. As the criteria of naturalness of conceptual construction and terminological choice may suggest, what counts is a combination of both. As regards naturalness of conceptual construction, suffice it to say that for imagination to be fruitful its stimulation has to be somehow regimented.

<sup>14</sup> A similar view seems to have been expressed in Yablo 2005: 98-100.

A certain kind of make-believe game, Walton says, can be ‘useful for articulating, remembering, and communicating facts’ about aspects of the game-independent world. He might have added that make-believe games can make it easier to reason about such facts, to systematize them, to visualize them, to spot connections with other facts, and to evaluate potential lines of research.

Following Walton (1993: 65), we can thus distinguish between ‘content-oriented’ and ‘prop-oriented’ make-believe.<sup>15</sup>

Which parts of the scientific discourse have to be considered as manifesting cases of living in the worlds of make-believe depends not only on ontological commitments but also on existential beliefs of the participants. If they believe in the existence of the given posits, there will be no make-believe. There are large parts of scientific discourse that do not manifest living in a world of make-believe. Thus there are parts of scientific discourse with reference to which the question of what sort of make-believe is instantiated does not arise. As for the remainder, observe that living in a world of make-believe in science is never an end for itself. It follows that it is never a content-oriented make-believe. There is a suggestion to the contrary in (Yablo 2005). He considers pure mathematics. He does not maintain without reservations that make-believe in pure mathematics is partly content-oriented (let alone that it is content-oriented *in toto*). But he considers it as a legitimate option on the grounds that:

pure mathematicians spend most of their time trying to work out what is true according to this or that mathematical theory [...and that:] the mathematician’s interest in working out what is true-according-to-the-theory is by and large independent of whether the theory is thought to be *really* true — true in the sense of correctly describing a realm of independently constituted mathematical objects (2005: 98-99).<sup>16</sup>

However, the independence referred to in the second part of the quotation seems to be due not to lack of interest in what is real, but to acknowledging (openly or somehow in the background) that mathematical theories need not be true *simpliciter* — that it is enough for them to be true by representation. What pure mathematicians are trying to do is mainly to check their conjectures against theories. Granted, checking conjectures against theories consists for the most part in working out what

<sup>15</sup> In the paper I refer the reader to, genuine cases of prop-oriented make-believe are discussed together with some related phenomena, like “recognition of the possibility of make-believe” (Walton 1993: 66-69; the quoted phrase is on p. 67). Generally speaking, the author is not very precise about what these phenomena are like. There are, perhaps, among them phenomena that are in various ways relevant to the present purposes. It brings about prospects for fruitful elaboration of the account of make-believe in science that is being advanced here. One of the guiding ideas is that to draw attention to the phenomena at issue makes it more likely that in some circumstances our consciousness fluctuates between states of make-believe and some other states.

<sup>16</sup> A note appended to the first part of the quotation explains what is meant by a theory (a very broad interpretation is intended). The second part has been supplemented by a note to the effect that “The intended contrast is with true-according-to-some-other-theory”.

is true according to them. But what makes the conjectures interesting is not exhausted by being true according to the given theories.

In denying that make-believe in science can be content-oriented, we are not thereby committed to the view that all make-believe in science is prop-oriented. It can be neither. It can be ‘hidden reality-oriented’.

## 10. THE LOGIC OF THEORETICAL FICTION

### 10.1. Susceptibility to logical regimentation

For a theoretical fiction to be useful, it is required that the theoretically relevant use of the linguistic devices which are involved in its positing — by way of reference relationships — be governed by clear logical rules. We may call such rules the ‘logic’ of the given *linguistic devices* (for the motivation of the scare quotes, cf. section 10.4), but it is often convenient to speak of the logic of the given *fiction* instead. The term “logic of theoretical fiction” can be used accordingly for the sake of generalization.<sup>17</sup> The logic of theoretical fiction as understood here is not to be confused with the branch of logic that assumes non-standard truth conditions purported to be distinctive of the discourse on non-being.

The stress with which the term “logic of theoretical fiction” is used has to be well balanced. No more stress on “fiction” than on “logic”. That the *logic* of theoretical fiction is the logic of theoretical *fiction* does not make it less committing, since the fiction we posit has to be real by representation. A useful fiction is, so to say, what we have made it to be, according to our theoretical aims. But there is no threat of arbitrariness. The theoretical aims and reality establish together strong constraints on what fictions can usefully be posited.

### 10.2. The positive and negative component

Whenever ontological commitment to objects believed to be fictitious is relevant to representation (can fruitfully be involved in representation), some sentences that ontologically commit to these objects are relevant too. The logic of theoretical fiction has to comprise rules that distinguish them from the rest. We must know such rules (both ‘knowing how’ and ‘knowing that’ may be involved). Thus the logic of theoretical fiction must have its positive and its negative component. On the positive side, we have, of course, the logic of the full-fledged theories that posit the given fictitious objects. It often happens, however, that such theories do not cover the whole area of fruitful application of the corresponding concepts.<sup>18</sup> Moreover, part of this area may

<sup>17</sup> A similar term has incidentally been used in Suárez 2009a: 5: “Vaihinger [...] advances what he takes to be a general *logic of scientific fictions*” (italics mine).

<sup>18</sup> When the limits of such area are determined as limits of what is expressible in a certain lan-

be occupied by provisional theoretical constructions with various degrees of maturity. If the systematization on the positive side is broad enough to cover all the demands, there is no need for rules on the negative side except for the following general rule: what is not allowed is forbidden.

Anyway, there is no symmetry. It will never happen that there is some elaboration on the negative side that allows to put things the other way round: what is not forbidden is allowed. The reason is that we never can be sure that enough has been forbidden. Granted, we could be almost sure, if the inconsistencies that can be revealed were necessarily confined to those accountable in terms of what we have recognized as a fiction. But inconsistencies in our theoretical constructions can always crop out whether or not theoretical fiction is involved. This asymmetry, observe further, is reminiscent of the asymmetry in logic between the logic of affirming and the logic of denying (more exactly: the logic of what deserves to be affirmed and what deserves to be denied).

### 10.3. Grammatical and representational meaningfulness<sup>19</sup>

What has to be accepted according to the logic of theoretical fiction is determined by the condition of, roughly, participation in representation. Now, speaking of representation gives often occasion for fruitful metaphorical use of semantic terms, such as “meaning”, “interpretation”, and the like. We can pursue the analogy on which these metaphorical use is based, drawing the above distinction in terms of ‘meaningfulness’ and ‘meaninglessness’. When there is a danger of misunderstandings we can distinguish between ‘grammatical’ and ‘representational’ meaningfulness (meaninglessness). Thus, what has to be accepted according to the logic of theoretical fiction can be said to be restricted to what is (representationally) meaningful.<sup>20</sup>

guage, in some cases no axiomatic theory can cover the whole area (due to limitations indicated by Gödel’s theorem of incompleteness). However, the positive component can be designed so as to include some criteria of evaluation of new axioms and/or new rules of inference.

<sup>19</sup> My account of the distinction between grammatical meaningfulness and representational meaningfulness owes much to helpful comments by Joanna Odrowąż-Sypniewska.

<sup>20</sup> This meaningful/meaningless distinction applies in all cases of representation, independently of whether they involve a theoretical fiction that mediates (cf. section 1.2) between language and reality. Thus, independently of the ontological status of numbers, in the case of ordinal measurement scales (e.g. scales for measuring hardness) it is meaningless (relative to the given numerical representation) to characterize the numerical values in terms of one being two, three, etc. times as great as another. Granted, in characterizing these numerical values in this way we do not break the rules of grammar, and we can formulate true sentences (if numbers are not fictitious objects). But the arithmetical facts that we thus speak of do not participate in representation. An additional point to be made here, to avoid misunderstandings, is that, *if* it is also inappropriate to characterize some *substances* as two, three, etc. times as hard as some other ones, it is due to a misconception of reality. Therefore what is conveyed by such characterization should not be couched in terms of lack of participation in representation. It should be couched in terms of representation based on misconception.

What can be meaningless in the above metaphorical sense includes questions. A question can have no meaningful direct answer. It may be instructive to quote Kripke in this connection (1980: 18):

Certainly the philosopher of ‘possible worlds’ must take care that his technical apparatus not push him to ask questions whose meaningfulness is not supported by our original intuitions of possibility that gave the apparatus its point.

Now, it would be illegitimate to divide questions into *grammatically* meaningful (grammatical) and *grammatically* meaningless (ungrammatical). There are no ungrammatical *questions*. What is more, there are no ungrammatical *sentences* in the interrogative. Granted, there are sentence-like sequences of words that feign grammaticality. Admittedly, such sentence-like word sequences can bring about confusion (usually similar to the one they manifest). But it is not so easy to utter ungrammatical sequences of words that purport to say something. It is perhaps easier to overlook hidden meanings. Anyway, violation of the logic of theoretical fiction by asking illegitimate (representationally meaningless) questions seems to be more harmful. But this is a matter of meaningfulness (meaninglessness) in the above metaphorical sense, and not of grammaticality. Since the sense of meaningfulness I recall seems to fit Kripke’s words, I do not intend my comment as a criticism but as an interpretation which brings the passage I quote close to what is at issue here; not without additional merit of gaining a suggestion on where to look for elucidation by examples. The suggestion points to the ontology and semantics of possible worlds as one of the sources.

#### **10.4. Logic and meaning. Is the term “logic of theoretical fiction” a misnomer?**

All linguistic means (expressions and otherwise) that contribute to expressing of what can be stated (asserted) have their logic. The logic of linguistics means is their contribution to truth conditions and entailment (or more generally: truth conditions of sentences relative to other sentences). Such contribution is part of their meaning. Now, that some sentences participate in representation in the spirit of fictionalism is not determined by the meaning of the linguistic devices that occur in them, which implies that the logic of theoretical fiction is not logic *tout court*. It might seem therefore that the term “logic of theoretical fiction” is a misnomer. However, all the rules of the logic of theoretical fiction govern the way the relevant linguistic means are used in cognition, which makes the incriminated use of the term “logic” not so artificial after all.

---

This kind of inappropriateness is not illustrative of what we are dealing here with.

## 11. THE MISGIVINGS ABOUT POSITING FICTION. THE ULTIMATE SOURCES AND POSSIBLE REMEDIES

### 11.1. Tarski's thesis on acceptability conditions reconsidered

Consider again Tarski's thesis on acceptability conditions. Observe that it is but quite recently that the spontaneous practice of positing fiction developed to the extent that it started to draw attention of philosophers (attention that is not marginal). That Tarski's thesis on acceptability conditions has a *prima facie* plausibility (cf. section 5) can be therefore looked upon as a disadvantageous heritage of the earlier stages of the development of science. This heritage is present in our thinking about the aims of science. Therefore it is also present in our thinking about the functions performed by theories. Whatever seems incompatible with it is likely to be viewed upon as incompatible with the aims of science, and thus intolerable.

Adherents of all philosophical standpoints look back in history to find precursors. Fictionalists are no exception. Now, in the case of fictionalism such search is not likely to reveal a glorious history.<sup>21</sup> But modern fictionalism is not a totally new standpoint, and we have to check whether the history of fictionalist ideas does not undermine the above argument from history. The history of astronomy is of special interest in this connection. Some Renaissance astronomers held fictionalist views with respect to the Copernican theory (Rosen 2005: 36-46). It can give rise to objections of the above sort. But to counter them it is perhaps enough to notice that the fictionalist solution we discuss originated in a very special social environment, under a hard pressure of religious dogma.

The way I appeal to history can also give rise to objections in connection with Bentham's theory of fictions. On Rosen's (2005: 46-56) account, it is dubious whether Bentham's theory of fictions involves genuine fictionalism (except for "a small commitment to a sort of fictionalism at the metalevel" (2005: 51)), the reason being that Bentham's paraphrases can be understood in a way which allows to regard them as bringing about an elimination of a sort. Whether or not Rosen's interpretation is correct is of course directly relevant to the present discussion. However, no matter whether it is correct or not, it draws attention to the following general relationship: the problem of theoretical fiction was not likely to get a permanent place on the philosophical agenda as long as superficial assessment of the prospects for elimination allowed philosophers to believe that all troublesome ontological commitment is easy to eliminate.

According to the above hypothesis concerning Tarski's thesis on acceptability conditions, its *prima facie* plausibility is a burden of history. There may be other burdens of history of the above origin among our views which give rise to objections against positing fiction. The claim that the given views are such burdens of history is

---

<sup>21</sup> Cf. Rosen 2005; see also section 1.3.



not necessary for undermining them. But it may be a significant contribution to argumentation.

### **11.2. Some possible prejudices about make-believe**

Recall the role of make-believe in scientific practice. Observe that there is a strong habit of thinking about make-believe as essentially non-cognitive, distinctive of the workings of the literary fiction, and of playing games. Such thinking may be looked upon as another burden of history. Be it as it may, it is likely to make us suspicious about make-believe in science. Thus it can make us suspicious about positing fiction. Granted, positing fiction does not make it necessary (in principle) to live in worlds of make-believe (cf. section 9). However, when efficiency is in the foreground (cf. section 9), it can easily occur to us that to let theories be built as they are and *not* to permit temporary living in worlds of make-believe is to defeat the purpose. Moreover, it is easy to notice that once positing fiction has been allowed, make-believe appears spontaneously, which makes it hardly avoidable.<sup>22</sup>

### **11.3. Truth and moral obligations**

The above sources of misgivings about positing fiction are of purely methodological character. Now, the notion of truth is also involved in moral obligations concerning the truth of what we say. Granted, it is rather unlikely that we let such moral obligations to interfere directly with our present-day scientific methodology. But it seems not implausible that as long as positing fiction remained, roughly, beyond philosophers attention, the above moral obligations illegitimately blended with methodological guidance. Of course, this factor should not be regarded as the strongest factor responsible for the misgivings about positing fiction. But it should not be totally neglected either.

### **11.4. The fear of insufficient logical regimentation**

The foregoing discussion has been confined to possible sources of misgivings about positing fiction that are directly connected with the aims of science. However, the possibility remains that the misgivings we consider are influenced also by some factors closer to pragmatic evaluations. Thus one can be inclined to think that positing objects believed to be fictitious is not sufficiently safeguarded against going

---

<sup>22</sup> Sainsbury makes a similar observation concerning make-believe in general, as contrasted with pretence: “Like belief, but unlike pretence, make-believe is often involuntary” (Sainsbury 2010: 12). The example that follows is not from science (“To open a novel with a receptive mind is to start make-believing”), but it does not make it irrelevant.

astray. That it does not admit of sufficient logical regimentation. Observe that what counts in this case is not positing fiction as such, but positing fiction beyond sufficient control. What is indicated as endangered is not our respect for truth but the efficiency of our cognitive activity. There are some passages in (Vaihinger 1924) that deserve a comment in this connection. On Vaihinger's view, the process of advancing hypotheses results in a tension due to a feeling of cognitive instability, or the like. Such tension may, as Vaihinger maintains, lead to illegitimate assertion (1924: 125). A similar tension, and an even stronger one, is claimed to arise in the case of positing fiction (1924: 126). It is enough, I think, to carefully characterize what is at issue here in order to avoid the misleading impression that it has something to do with the misgivings arising from the assessment of the extent to which positing fiction can evade our control. Note that the feeling of cognitive instability Vaihinger draws our attention to occurs both in the case of advancing hypotheses and in the case of positing fiction. The difference is in degree.

### **11.5. Some possible remedies**

The above list of ultimate sources of worries about positing fiction is not uniform in respect of what can bring about a remedy. As regards the possible burdens of history that have been indicated, in some cases it is enough (roughly) to draw attention to them to dispose of them. The same remedy applies perhaps to the case of Tarski's thesis on acceptability conditions, as well as to the case of possible interference with moral obligations. However, what has been said of the views on the role of make-believe gives rise to questions of remedy that cannot be answered in such a simple way. On the other hand, it can be maintained that some possible remedies have already been indicated in section 9. The account of belief imitation provided in this section makes it plain that living in worlds of make-believe in science is part and parcel of research processes involving belief imitation of various sorts, including its less debatable varieties, which may result in conferring to make-believe a bit of familiarity. To turn to the other sources of misgivings about positing fiction, we have to raise the question of whether it is possible to dispose of the fear of positing fiction beyond sufficient control. But once again a partial answer seems to be at hand thanks to the foregoing discussion. The moral to be drawn from the considerations of section 10 is that positing fiction is likely to admit of satisfactory logical regimentation.

## **12. ARE ELIMINATION PROGRAMMES A THREAT TO FICTIONALISM?**

Suppose an elimination programme (cf. section 7) has been developed that gives good prospects for fruitful elimination of all posits believed to be fictitious. To consider a certain elimination programme as so formidably powerful would not preclude

being a fictionalist, if positing fiction remained as a theoretical possibility. However, positing fiction being no more than a purely theoretical possibility would make fictionalists' position weaker since it could suggest that something goes wrong (that some flaws have been overlooked). Thus whatever can be regarded as a contribution to designing an elimination programme of the above sort is a threat to fictionalism, although not necessarily a very serious one. What, precisely, are such threats to fictionalism at present? The considerations to follow are intended as relating to this issue. They do not cover all relevant subject matter. On the contrary, there are important omissions.<sup>23</sup>

Observe that lack of interpretation deprives our derivations of the possibility of auxiliary testing for validity against logical intuitions that manifest our 'knowledge' of meanings. Moreover, it cuts us off from almost all sources of inspiration important for designing new theories. To turn to the non-instrumentalist elimination, we should perhaps note, in the first place, that all such elimination is parasitic on the theories that are intended to be eliminated.<sup>24</sup> To be more specific, advocates of non-instrumentalist elimination never build new theories from scratch. They modify the theories responsible for the ontological commitments they intend to eliminate. Moreover, in most cases elimination depends on the theory that is replaced not only at the stage of designing the new theory, but also at the stage of putting it to work. As for the first possibility, suppose somebody denies the existence of properties and relations while holding the view that predicative sentences ontologically commit to properties and relations denoted by the respective predicates.<sup>25</sup> Such ontological commitment can be eliminated by reformulation that replaces the given sentences by the corresponding Ramsey sentence. The ontological commitment that 'takes its place' need not be the ontological commitment to properties and relations in general (or properties and relations of some sort), which on the above assumptions would obviously be of no help. One can choose as values of variables involved in the operation in question abstract objects of some other category, e.g. sets. But for a nominalist it will not do either. Generally speaking, the method of using a Ramsey sentence seems to have a rather narrow range of applications. Moreover, there is an obvious decrease in respect of formal convenience.

In using a Ramsey sentence we would probably avail ourselves of the old theory only on the stage of designing the new one. The involvement of the old theory goes further in the case of applications of the method of 'axiomatization within a system'

---

<sup>23</sup> Thus I leave aside, e.g., the whole debate concerning the extent to which Field's demonstration-of-non-indispensability programme is successful. For Carnap's programme of *der logische Aufbau der Welt*, cf. section 5.

<sup>24</sup> In making this point, I draw on Bennett 2003: 156, Craig 1956: 50, Divers 1995: 86, and Rosen 1990: 332, 336, 337.

<sup>25</sup> On this view, sentences, e.g., "Snow is white" and "Snow has whiteness" alike ontologically commit to the property of whiteness.

due to Craig.<sup>26</sup> Suffice it to say that the decision procedure for the set of axioms of a theory built by using this method (it is an infinite set) involves knowledge of the structure of the initial theory (encoded by Gödel numbering). Neither is the dependence restricted to the stage of designing a theory in the case of Rosen's modal fictionalism.<sup>27</sup> To turn back to Craig's method of 'axiomatizability within a system', it is worth noting that it produces theories with axioms that do not satisfy the condition of bringing about elucidation of terms that occur in them (Craig 1956: 49, 52). Thus the elimination programmes that can be based on it can hardly be claimed to satisfy the functional adequacy condition (for the notion of functional adequacy, cf. section 7).

Let me close this section by some further remarks concerning modal fictionalism. Rosen's presentation of modal fictionalism starts with formulations intended as indicating the guiding ideas (or at least some of them). One of the ideas indicated with this intention is that ontological commitment to possible worlds can be eliminated by placing the given sentence in the scope of an operator ('story prefix') of the form "according to...", the place of the dots being occupied by a singular term designating some ontological theory that is or 'contains' a theory of possible worlds. As regards the criteria for choice of the theory to be used as an appropriate 'story', the guiding idea is conveyed by the following formulation:

The chief constraint [on the choice of the 'story'], of course, is that it be one according to which the usual claims about possible worlds — for example, the claim that there are blue swan worlds — are true (Rosen 1990: 332).

It may not be *prima facie* clear what is the rationale of this criterion for choice. Looking for an answer, we may notice that a story prefix chosen in conformity with the criterion that has been indicated forms true sentences with sentences expressing

---

<sup>26</sup> Besides the paper that has already been referred to, cf. Craig 1953. The later of these two papers provides an informal exposition of and a philosophical elaboration on the metatheoretical results presented in the earlier one. For comments on the dependence on the old theory, cf. especially Craig 1956: 50.

<sup>27</sup> Cf. especially Rosen 1990 and Rosen 1995. If I properly grasp what is intended, modal fictionalism aims at elimination. It is not a version of fictionalism, but a view that combines opposing modal realism with an elimination solution. Thus the term "modal fictionalism" seems to be a misnomer not only due to its first constituent (modal fictionalism is not concerned with modality in general) but also due to the second one. This term seems to have been coined as contrasting with the term "modal realism" introduced by David Lewis for the position he takes in the ontology of possible worlds. But fictionalism is a methodological standpoint, and not an ontological one. It bears on ontological issues but indirectly (cf. section 7). Of course, to the extent modal fictionalism is an ontological view, it is legitimate to contrast it with modal realism, but my remarks concern the terminology. Incidentally, it may be noted that Lewis was not content with the above terminological convention. In Lewis 1986 he apologized. And he tried to counteract some misunderstandings by indicating what modal realism is not (p. viii). To return to modal fictionalism, observe that as far as it involves an elimination solution, it can be contrasted with fictionalism as presented in Rosen 2005. As regards distinctively fictionalist ideas in possible worlds ontology, cf. Nolt 1986.

‘the usual claims about possible worlds’. But why should we care about such sentences (as contrasted with *other sentences about possible worlds*), if we do not believe in the existence of possible worlds? This question has not been raised, let alone answered, but the answer seems to be almost at hand (which precludes any serious criticism concerning the relevant details of presentation): the only reason why we may be interested in making the above distinction is that we regard sentences expressing the usual claims about possible worlds as conveying something important indirectly *via* representation.

The way the first of the above two guiding ideas has been conveyed gives rise to an interpretive difficulty too. Observe that the elimination of ontological commitment to possible worlds is not an aim in itself. It has to be an elimination that preserves something important — if not in all relevant theoretical contexts, then at least in some of them. What are the theoretical contexts that are regarded as allowing the above treatment? The answer comes in the wake of presentation of the second of the above two guiding ideas, the one concerned with the criteria for choice of the ‘story’.<sup>28</sup> Let *PW* denote, as in Rosen’s papers that have been referred to, the ‘story’ (theory) that has been initially chosen (i.e. has been chosen in Rosen 1990). Another choice has been considered subsequently as a possible alternative (Rosen 1995), but the difference does not matter for present purposes. As far as possible worlds ontology is concerned, *PW* is, loosely speaking, based on Lewis’s modal realism. Now, following Rosen we may consider the schema:

(1)  $P$  iff  $P^*$

where  $P$  is an arbitrary modal sentence, and  $P^*$  is “the modal realist’s non-modal paraphrase of  $P$  in the language of possible worlds” (1990: 335).<sup>29</sup> Suppose that the place of  $P$  is occupied by the sentence (Rosen’s example) “There might have been blue swans” (observe that this sentence expresses a theorem of our common sense theory of the modal order of the world). The conjunction with the respective equivalence ontologically commits to possible worlds in which there are blue swans. This ontological commitment is eliminated when (1) is replaced by the schema:

(2)  $P$  iff, according to *PW*,  $P^*$

As this replacement indicates, the theoretical contexts of elimination Rosen’s method of elimination has been devised for are those we have just discussed, i.e. the contexts of rules for providing ordinary modal sentences with a paraphrase in the

---

<sup>28</sup> There are also some interpretive hints earlier (see especially Rosen 1990: 327) but not sufficiently informative if taken separately.

<sup>29</sup> To characterize the paraphrases in the language of possible worlds as non-modal is objectionable (modal realism does not make such characterization acceptable without reservations), but whatever the objections and the ways Rosen could answer them, they have no bearing on the issue that is our concern here.

language of possible worlds (cf. Rosen 1990: 327 for additional evidence).<sup>30</sup> Accordingly, at some places (Rosen 1990: 332, 336, 340, 351), the view seems to have been expressed that the schema (2) gives an adequate rule for *paraphrasing* that replaces the one given by (1). At the very end of the paper it turns out that Rosen is ready to diagnose the relationship at issue as somewhat weaker. But he seems to believe anyway that the equivalences of the form (2) are necessary sentences (1990: 364).

The above brief presentation of the basic tenets of modal fictionalism indicates that what counts for the modal fictionalist in her choice of a ‘story prefix’ is not only the ontological ‘story’ itself, but also an accompanying choice of some rules for paraphrasing, Lewis’s views having been chosen as the views to draw on due to both the content of his possible worlds ontology and the way he uses its language to paraphrase ordinary modal sentences. But what are the criteria for choice of such rules? It is tacitly assumed, one can guess, that our usual claims about possible worlds reflect, on purely conceptual grounds, our ordinary modal intuitions (e.g. the claim that there are blue swan worlds reflects the modal intuition that there might have been blue swans)<sup>31</sup>. Now, this relationship determines certain rules of paraphrasing. Call them ‘reflecting rules’. If the reflecting rules are followed, ordinary modal sentences that are true according to the ordinary modal intuitions receive a paraphrase that is true according to the usual claims about possible worlds, e.g. the sentence “There might have been blue swans” receives the paraphrase “There are blue swan worlds”.

Taking the schema (1) we can put it this way: if Lewis’s rules for paraphrasing are reflecting rules, i.e. the asterisk operation is defined indirectly in terms of such rules, the schema (1) correlates sentences that are true according to ordinary modal intuitions with sentences that are true according to the usual claims about possible worlds. Now, if  $PW$  respects the usual claims about possible worlds, sentences that are true according to them form true sentences with the operator “according to  $PW$ ”. Finally, we get therefore: if  $PW$  respects the usual claims about possible worlds, and Lewis’s rules for paraphrasing are reflecting rules, i.e. the asterisk operation is defined indirectly in terms of such rules, the schema (2) correlates sentences that are true according to ordinary modal intuitions with true sentences of the form “According to  $PW$ ,  $p$ ”.

Is Rosen’s elimination programme functionally adequate? Consider first the aim of providing paraphrases for sentences of the ordinary modal idioms. As has been noted above, Rosen himself casts doubt on the claim that sentences on the right hand side of equivalences of the form (2) can be treated as paraphrases of sentences on the left hand side. There is also a more severe criticism due to Divers: the basic idea is

---

<sup>30</sup> What is indicated here are, of course, only the direct theoretical contexts. The paraphrases serve as a bridge between the ordinary modal intuitions and the wide theoretical framework of the ontology of possible worlds (cf. Rosen 1990: 330).

<sup>31</sup> This guess seems very natural in view of the above comment on how the usual claims about possible worlds are involved in Rosen’s criterion of choice of the ‘story’.

that it is strongly counterintuitive to maintain that the *metatheoretical* sentences of the form “According to  $PW$ ,  $p$ ” can be paraphrases of ordinary modal sentences (1995: 83-84). It has to be observed, further, that sentences on the right hand side of the equivalences of the form (2) cannot be used to give truth conditions of sentences on the left if it has to be done in conformity with standard semantic methods (Divers 1995).

One can perhaps make the following tentative generalization: the form of the sentences on the right hand side seems to exclude a contribution to a fruitful representation of the modal facts expressible in ordinary modal idioms, no matter what mode of representation has been chosen. A closely related objection can also be raised in connection with Rosen’s choice of the possible worlds ‘story’ (it bears directly on the links with the broader theoretical background). There are rival ontological theories of possible worlds that deserve consideration from Rosen’s perspective, and it may be maintained that there are no good criteria for choosing among them (Sainsbury 2010: 190-192). The upshot is that our question concerning functional adequacy seems to require an answer in the negative: the condition of functional adequacy is not met.<sup>32</sup>

### 13. FINAL REMARKS. AIMING AT TRUTH

Recall the methodological maxim “Science aims at truth” (the maxim of aiming at truth). In section 1 its reliability has been challenged. The above defence of fictionalism (more exactly: its version that countenances make-believe in science) bears on this issue. The maxim of aiming at truth is in need of interpretation. According to one of the interpretations that are plausible enough to be worthy of mention, the maxim of aiming at truth demands that theorems of scientific theories be true or ‘true by representation’. Of course, this demand is perfectly compatible with fictionalism: a fictionalist is not tempted to be *more* tolerant towards theoretical fiction. But in the light of what has been said above, the maxim of aiming at truth can be accused of being *misleading*. The interpretations of ‘aiming at truth’ that are most likely to be *tacitly* assumed make it closely akin to Tarski’s thesis on acceptability conditions. Now, it has turned out that we have good reasons to abjure this thesis as a burden of history.

---

<sup>32</sup> Some tenets of modal fictionalism seem to have a rationale conferred to them outside the confines of possible solutions to the problem of theoretical fiction. The point is that modal fictionalism can be viewed upon as advancing a positive appraisal of the ontology of possible worlds (roughly its Lewisian version) in respect of its capacity for representing modal facts expressible in ordinary modal idioms.

## REFERENCES

- Ajdkiewicz K. (1956/1978), *Conditional Statement and Material Implication* [in:] *The Scientific World Perspective and Other Essays*, J. Giedymin (ed.), Dordrecht: D. Reidel, 222-238 (transl. from the Polish original).
- Barberousse A., Ludwig P. (2009), *Models as Fictions* [in:] Suárez 2009b: 56-73.
- Bayne T., Fernández J. (2009), *Delusion and Self-Deception. Mapping the Terrain* [in:] *Delusion and Self-Deception*, T. Bayne, J. Fernández (eds.), New York: Psychology Press, 1-21.
- Bennett J. (2003), *A Philosophical Guide to Conditionals*, Oxford: Oxford University Press.
- Bird A. (2007), *Nature's Metaphysics*, Oxford: Oxford University Press.
- Bokulich A. (2009), *Explanatory Fictions* [in:] Suárez 2009b: 91-109.
- Brykczyński P. (2009), *Mereology and the Concept of Sum* [in:] *Logic, Methodology and Philosophy of Science at Warsaw University 5*, A. Horecka (ed.), Warszawa: Wydawnictwo Naukowe Semper.
- Chisholm R. M. (1981), *First Person*, Brighton: Harvester Press.
- Church A. (1958), *Ontological Commitment*, „Journal of Philosophy” 55(23), 1008-1014.
- Corry L. (2004), *Modern Algebra and the Rise of Mathematical Structures*, Basel: Birkhäuser.
- Craig W. (1953), *On Axiomatizability within a System*, „Journal of Symbolic Logic” 18(1), 30-32.
- Craig W. (1956), *Replacement of Auxiliary Expressions*, „Philosophical Review” 45(1), 38-55.
- Divers J. (1995), *Modal Fictionalism Cannot Deliver Possible Worlds Semantics*, „Analysis” 55(2), 81-89.
- Elgin C. Z. (2009), *Exemplification, Idealization, and Scientific Understanding* [in:] Suárez 2009b: 77-90.
- Field H. H. (1980), *Science without Numbers*, Princeton (NJ): Princeton University Press.
- Field H. H. (1989), *Realism, Mathematics and Modality*, Oxford: Blackwell.
- Fine A. (1993/2009), *Fictionalism*, „Midwest Studies in Philosophy” 18; reprinted in Suárez 2009b: 19-36.
- Fraenkel A. A., Bar-Hillel Y., Levy A. (1973), *Foundations of Set Theory*, Amsterdam: North-Holland.
- Hempel C. G. (1965), *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: Free Press.
- Kalderon M. E. (ed.) (2005), *Fictionalism in Metaphysics*, Oxford: Oxford University Press.
- Kant I. (1998), *Critique of Pure Reason*, Cambridge: Cambridge University Press.
- Kripke S. (1980), *Naming and Necessity*, Oxford: Basil Blackwell.
- Lewis D. (1986), *On the Plurality of Worlds*, Oxford: Basil Blackwell.
- Loux M. J. (2006), *Metaphysics*, New York: Routledge.
- Meinong A. (1971), *Über Gegenstandstheorie* [in:] *Alexius Meinong Gesamtausgabe*, vol. 2, R. Haller, R. Kindinger, R. Chisholm (eds.), Graz: Akademische Druck- u. Verlagsanstalt, 481-530.
- Meinong A. (1977), *Über Annahmen* [in:] *Alexius Meinong Gesamtausgabe*, vol. 4, R. Haller, R. Kindinger, R. Chisholm (eds.), Graz: Akademische Druck- u. Verlagsanstalt.
- Nolt J. E. (1986), *What Are Possible Worlds?*, „Mind” 95(380), 432-445.
- Orenstein A. (1980), *What Makes Substitutional Quantification Different?* [in:] *Sprache, Logik und Philosophie. Akten des vierten internationalen Wittgenstein Symposiums*, R. Haller, W. Grassl (eds.), Vienna: Hölder-Pichler-Tempsky, 346-349.
- Popper K. R. (1979), *Objective Knowledge. An Evolutionary Approach*, Oxford: Oxford University Press.
- Quine W. V. O. (1960), *Word and Object*, Cambridge (MA): MIT Press.



- Quine W. V. O. (1980/1961a), *Two Dogmas of Empiricism* [in:] *From a Logical Point of View*, Cambridge (MA): Harvard University Press, 20-46.
- Quine W. V. O. (1980/1961b), *Logic and the Reification of Universals* [in:] *From a Logical Point of View*, Cambridge (MA): Harvard University Press, 102-129.
- Quine W. V. O. (1966), *The Ways of Paradox and Other Essays*, New York: Random House.
- Quine, W. V. O. (1966a), *On Mental Entities* [in:] Quine 1966: 208-214.
- Quine W. V. O. (1966b), *The Scope and Language of Science* [in:] Quine 1966: 215-232.
- Quine W. V. O. (1966c), *Posits and Reality* [in:] Quine 1966: 233-241.
- Quine W. V. O. (1969), *Epistemology Naturalized* [in:] *Ontological Relativity and Other Essays*, New York: Columbia University Press, 69-90.
- Quine W. V. O. (1970), *Philosophy of Logic*, Englewood Cliffs (NJ): Prentice Hall.
- Quine W. V. O. (1977), *The Ways of Paradox and Other Essays*, revised and enlarged edition, Cambridge (MA): Harvard University Press.
- Quine W. V. O. (1977a), *The Limits of Knowledge* [in:] Quine 1977: 59-67.
- Quine W. V. O. (1977b), *On Multiplying Entities* [in:] Quine 1977: 259-264.
- Quine W. V. O. (1981), *Things and Their Place in Theories* [in:] *Theories and Things*, Cambridge (MA): Harvard University Press, 1-23.
- Rosen G. (1990), *Modal Fictionalism*, „Mind” 99(395), 327-354.
- Rosen G. (1995), *Modal Fictionalism Fixed*, „Analysis” 55(2), 67-73.
- Rosen G. (2005), *Problems in the History of Fictionalism* [in:] Kalderon 2005: 14-64.
- Rosen G., Dorr C. (2002), *Composition as a Fiction* [in:] *The Blackwell Guide to Metaphysics*, R. M. Gale (ed.), Oxford: Blackwell, 151-174.
- Sainsbury R. M. (2010), *Fiction and Fictionalism*, London: Routledge.
- Searl J. R. (1992), *The Rediscovery of the Mind*, Cambridge (MA): MIT Press.
- Shapiro S. (2007), *Philosophy of Mathematics and Its Logic. Introduction* [in:] *The Oxford Handbook of Philosophy of Mathematics*, S. Shapiro (ed.), Oxford: Oxford University Press, 3-28.
- Suárez M. (2009a), *Fictions in Scientific Practice* [in:] Suárez 2009b: 3-15.
- Suárez M. (ed.) (2009b), *Fictions in Science*, New York–London: Routledge.
- Tarski A. (1944), *The Semantic Conception of Truth and the Foundations of Semantics*, „Philosophy and Phenomenological Research” 4(3), 341-375.
- Teller P. (2009), *Fictions, Fictionalization and Truth in Science* [in:] Suárez 2009b: 235-247.
- Vaihinger H. (1924), *The Philosophy of “As If”*, London: Routledge.
- Vaihinger H. (1927), *Die Philosophie des Als Ob*, Leipzig: Felix Meiner.
- Van Fraassen B. C. (1980), *The Scientific Image*, Oxford: Oxford University Press.
- Walton K. L. (1993), *Metaphor and Prop Oriented Make-Believe* [in:] Kalderon 2005: 65-87.
- Wittgenstein L. (2001), *Philosophical Investigations*, Oxford: Blackwell.
- Yablo S. (2005), *The Myth of the Seven* [in:] Kalderon 2005: 88-115.