Routledge
Taylor & Francis Group

Philosophical Psychology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/cphp20

The land of make-believe: metaphor, explanation, and fiction in Toon's psychological world

Tamás Demeter, László Kocsis & Krisztián Pete

To cite this article: Tamás Demeter, László Kocsis & Krisztián Pete (06 Dec 2024): The land of make-believe: metaphor, explanation, and fiction in Toon's psychological world, Philosophical Psychology, DOI: 10.1080/09515089.2024.2436527

To link to this article: https://doi.org/10.1080/09515089.2024.2436527

PHILOSOPHICAL PSYCHOLOGY
https://doi.org/10.1080/09515089.2024.2436527

Routledge
Taylor & Francis Group

The land of make-believe: metaphor, explanation, and fiction in Toon's psychological world

Tamás Demeter[a], László Kocsis[b] and Krisztián Pete[b]

[a]ISPS, Corvinus University of Budapest, and Institute of Philosophy, HUN-REN Research Centre for the Humanities, Budapest, Hungary; [b]Department of Philosophy, University of Pécs, Pécs, Hungary

ABSTRACT

In *Mind as Metaphor*, Adam Toon interprets folk psychological discourse metaphorically. Based on Kendall Walton's theory of metaphor, he argues that folk psychology ought to be understood in terms of prop-oriented make-believe that relies on representationally essential metaphors. Toon insists that this fictionalist view of everyday mental talk preserves what we commonly think folk psychology can achieve: it does not only rationalize but explains behavior causally. In this paper, first we raise concerns about Toon's characterization of folk psychology as metaphorical. Then we proceed to show that representationally essential metaphors are incompatible with genuine causal explanation. And finally, we argue that no construal of mental fictionalism can preserve the epistemic virtues we commonly ascribe to folk psychological discourse.

ARTICLE HISTORY
Received 4 September 2024
Accepted 26 November 2024

KEYWORDS
Mental fictionalism; metaphor; explanation; make-believe

1. Introduction

According to the most widespread view, folk psychology (FP) is constitutive in explaining, predicting, and understanding human behavior. For many, FP can play this role because it succeeds in describing our mental architecture, and so it can track the causal background of behavior. But for some others, FP statements cannot be literally true because FP terms fail to refer. Adam Toon aims to reconcile these two claims in his book, *Mind as Metaphor*. He suggests that while ascribing thoughts, beliefs, desires, we only *pretend* that these entities exist by using metaphors that resist literal paraphrase. Despite being a figurative discourse, FP can be epistemically valuable: a fictionalist about FP can "allow that, even if they don't describe inner causes, folk psychological explanations *are* genuine causal explanations" (p. 52). Thus, Toon's fictionalism rests on two crucial claims: i) FP operates with *representationally essential metaphors* (REM) resisting literal paraphrases, in terms of which we provide ii) *genuine causal explanations* (GCE).

CONTACT Tamás Demeter tsd2333@gmail.com ISPS Corvinus University of Budapest, Budapest, Fővám tér 8 1093, Hungary

But we feel we are left in the dark as to how such a discourse is possible at all. In this paper we argue that Toon's attempt to take FP a causally explanatory discourse that operates with REMs leads to a dilemma without an apparent solution. Either we insist that in terms of FP we can provide GCEs and so identify the causes of behavior, but then FP cannot operate with REMs. Or we insist that FP relies on such metaphors, but then we cannot provide GCEs in terms of FP.

Having sketched Toon's position (sect. 2), we will argue that his "pretence fictionalism" faces at least three substantive challenges. First (sect. 3), we run a "test for nonliteralness" (that of Rosen & Burgess, 2005) on FP. We show that Toon's construal fails to pass, and therefore it cannot be considered a metaphorical discourse. Secondly (sect. 4.), we go on to explore if REMs can be reconciled with causal explanatory power and will raise doubts about the prospects of reconciliation. And finally (sect. 5), we question whether Toon's account is indeed a form of fictionalism, and we point out an alternative more fitting to reasonable criteria.

2. Toon's attempt to have a cake and eat it too

As Toon points out, there are at least two "conflicting intuitions over folk psychological explanations of behavior." (p. 51) First, in answering questions about why people behave the way they do, "it is natural to think that we are offering a causal explanation" (p. 49). Second, the relation of mental states and behavior seems not only causal but also conceptual: "it is not only a contingent matter that the desire for ice cream tends to lead to eating ice cream" (p. 50). That is, behavior and its supposed mental causes are conceptually linked.

According to Toon, there are two opposing views about folk psychological discourse based on these intuitions: Cartesianism and anti-Cartesianism. For Toon's Cartesians, FP is not just an intuitive way of talking but a sort of (proto-scientific) theory about mental states as inner representations and about their role in explaining and predicting behavior. Cartesians hold that FP aspires to be true and to give GCEs. However, Cartesians cannot rescue the intuition that mental states seem to be conceptually and not only causally related to the behavior explained by them. Toon's anti-Cartesians, like behaviorists and instrumentalists, admit that there are conceptual connections between behavior and the mental states considered as its causes. They argue that FP explanations are not GCEs "but instead serve to fit someone's behavior into a larger pattern" (p. 51). However, anti-Cartesians cannot effectively accommodate the strong intuition that mental states are causally responsible for behavior.

Toon agrees with the Cartesians that we causally explain behavior in terms of inner states; *and* he agrees with the anti-Cartesians that in doing so

we do not refer to those inner states. He sees as one of the great advantages of his fictionalism that it can save both intuitions and avoid the problems of both the Cartesians (how in terms of mental states we can explain behavior) and the anti-Cartesians (why mental states do not cause behavior).

It is for this reason that invoking metaphors seems to be so valuable:

> one of the main reasons that we value metaphors is that they expand our powers of expression, allowing us to say things that we could not say without them. The same lesson applies to the metaphors that we use to describe the mind. These metaphors also add to our powers of expression, allowing us to assert things about people that we could not say otherwise. (p. 35)

Toon gives a Waltonian twist to this view (Walton, 1990): when we explain behavior by mental states, we participate in a game of prop-oriented make-believe. People and their behavior are the *props* in this game, and we describe them in terms of mental states, *as if* people had these states (p. 23). His stance is aptly summarized when he says that "our folk psychological concepts are metaphorical and governed by games of pretence" (p. 55) – which is to say that we deploy these concepts to talk about props (people and their behavior) in ways that they are (literally) not. The metaphors in which we talk about an inner mental world project the realm of public representations, the world of notebooks (beliefs), courtrooms (judgments), committees (decisions), onto the causal background of people's behavior: *as if* it were in the way these public realms literally are.

3. A test for nonliteralness

Influenced by Walton (1993) and Yablo (1998), Toon introduces the idea of a REM, a peculiar kind of metaphor: they express a truth that cannot be expressed otherwise, and for this reason, they do not have literal paraphrase (pp. 34–35). Toon characterizes FP terms as REMs whose criteria can be seen as follows (see Grant, 2013, p. 128):

(i) A REM must represent its object as having a certain property.
(ii) This property of the object makes the use of metaphor appropriate.
(iii) The property cannot be picked out literally.

Toon borrows his favorite metaphor of "angry clouds" from Yablo (1998, p. 250). The lesson he draws here is this. We cannot say exactly which property storm clouds that we call angry have in common. And similarly, we cannot say what all the kinds of behavior on which an FP ascription rests have in common.

> For the fictionalist, the failure of the behaviorist project suggests that the metaphors of folk psychology are representationally essential. If we say Mark believes the No. 73 bus

> goes to Oxford Street, we are saying that he is in some state *S*. Which state? Well, the state such that it is appropriate to pretend in the way that we do. There might be no alternative way of picking out this state, apart from by means of our pretense. In particular, there might be no way of specifying this state in terms of behavior. (. . .) The problems faced by behaviorism thus turn out to be an instance of a more general phenomenon, namely the difficulty of giving a literal paraphrase for metaphors and figurative language. (p. 35)

But if FP contains just metaphors in a game of pretense, how is it that when competent speakers use them, they think they are using them *literally*? We do not deny that we "usually pay little or no attention to the fact that we are speaking metaphorically" (p. 25), and that we sometimes have difficulty recognizing that we use metaphors. But a metaphorical way of speaking is always distinguishable from a literal one, even if we need help seeing the distinction. One way of seeing this is to recognize those cases when we are misusing metaphors and take them in a literal sense. We doubt that one can make this mistake in everyday mental talk.

Let's run a test for nonliteralness based on Rosen and Burgess (2005), which was originally advanced against Yablo's mathematical fictionalism:

> We submit that whenever a bit of language is used nonliterally, it is possible for an interlocutor to misconstrue it by taking it literally, and for the competent speaker to recognize this misunderstanding and correct it by pointing out that the remark was not meant literally. Certainly in all *clear* cases of figurative language—and it is worth stressing that the boundary between figurative and literal is as fuzzy as can be—the nonliteral character of the linguistic performance will be *perfectly obvious* as soon as the speaker is forced to turn attention to the question of whether the remark was meant literally. (Rosen & Burgess, 2005, p. 533)

Consider cases where competent speakers correct those who misconstrue metaphors for literal expressions. If we say Crotone is on the arch of the Italian boot and someone (a child, a foreigner, or an incompetent speaker) replies that "I never thought that a town could fit on the sole of a boot", then we can tell them that it was just a metaphor, and it should not be taken literally. If we say that angry clouds are coming and someone else asks, "Who are the clouds angry at?", then again, we can tell this person that it was just a metaphor. Literally speaking, the clouds are not angry at anyone.

But if we say that it was Mark's belief that caused him to reply, "the No. 73 goes to Oxford Street", and someone tells us that "Stop taking this metaphor literally", we would be puzzled about how to stop doing so. We could not even come up with a question here suggesting a literal misconstruction. But we invite the reader to indulge in the exercise. Notice that silly questions can be asked about literal statements (for example and characteristically about entire systems like organisms, computers, and universities),[1] but these questions only indicate a category mistake rather than

a metaphor. To put it succinctly: while many metaphors are category mistakes, category mistakes are not necessarily metaphors. To infer from a category mistake to the presence of metaphor would commit the fallacy of affirming the consequent.[2]

Moreover, Toon's silly questions seem to be motivated by philosophical commitments independent of the inclinations of competent speakers to construe FP ascriptions literally. For example, to ask silly questions about the location or shape of mental states, we must commit that mental states have physical properties. But why do we assume that competent FP speakers have such robust philosophical commitments? There are indeed philosophical approaches, e.g., those Toon calls Cartesian, which claim that FP is a sort of (proto)theory according to which beliefs and desires are of some nature, physical or nonphysical. However, taking FP literally does not entail such commitments, only commitments to the causal role of FP states. Toon ascribes more philosophical refinement to FP speakers than this, and he conflates literalness with taking robust philosophical commitments (see Burgess, 2004, p. 25). A competent speaker does not need to be a member of a philosophical sect to take FP statements literally.

We do not doubt that sometimes using metaphors is the best way to describe something, nor that sometimes we are unaware that we are describing something metaphorically. What we do doubt is that we could misconstrue metaphors systematically by taking them literally, without even being able to realize the misconstruction. We agree with Rosen and Burgess: "if there can be no literalistic misconstrual, then the language was not figurative in the first place" (2005, p. 533). If there is no place for mistaking FP claims for literal claims, and so they cannot pass the test for nonliteralness, then we have a good reason for not taking them to be parts of a metaphorical discourse. They must mean what they literally mean.

4. REMs *and/or* GCEs

Let's put the test for nonliteralness on one side and assume for the sake of argument that Toon is right: the correct interpretation of folk psychological discourse is the kind of fictionalism he proposes. Accordingly, everyday belief-desire statements are metaphorical moves in a game of make-believe. Predicting and explaining are parts of the game. When we explain Mark's statement, we are not asserting that there is a mental state, a belief that "the No. 73 goes to Oxford Street" which causes him to say what he said, but "we are simply pretending that this is the case" (p. 52). But can one give, *systematically*, GCEs while pretending to give GCEs?[3]

The pretense fictionalist must first address the challenge that "we do not ordinarily feel as if we are speaking metaphorically, much less pretending, when we say what people think or want or feel" (p. 25).

For Toon, when we *pretend* that we have mental states, we also make *genuine* assertions. These are genuine assertions about real people and their real behavior, but we pretend that they are in a way that they are actually not. We make these assertions without noticing that we are using metaphors: "often, we do not notice when we are speaking metaphorically, especially when metaphors are familiar to us" (p. 25). And this is Toon's response to the obvious phenomenological objection that we do not feel like using metaphors in making FP ascriptions.

Toon's concept of metaphor is undifferentiated and vague. It makes a difference what kind of metaphor we take FP statements to rely on. To show this, let's distinguish between three basic kinds with different characteristics:

(i) dead metaphors;
(ii) idiomatic or standardized metaphors;
(iii) poetic (creative) metaphors.

These different kinds of metaphor face the tribunal of the above criteria for REM with different consequences. To put it succinctly: categories i) and ii) can be GCEs precisely because they are not REMs; category iii) are REMs but for this reason they cannot be GCEs.

Let's turn to i) and ii) first. Yablo's (1998, p. 250) examples that inspire Toon are problematic as REMs because they are either dead metaphors, like "computer viruses" and "basins", or idiomatic metaphors, like "angry clouds". *Pace* Yablo, these metaphors do not meet his *third* criterion of non-paraphrasability (see the beginning of Section 3), because they are either literal expressions or can be paraphrased. Toon's list of metaphors also comprises of dead, idiomatic, and highly standardized metaphors (pp. 25–26). Dead metaphors are the most conformable to Toon's idea of unnoticed metaphors – but as he rightly acknowledges (pp. 40–41), dead metaphors are not proper metaphors, but literal expressions. The statements in which dead metaphors occur are straightforward not only in *what* they assert, but also in *how* they assert it.

More importantly, while metaphors are widely taken to be category mistakes (also by Toon, p. 26), dead metaphors are not. Other metaphors are category mistakes, and thus we can misconstrue them literally and it feels kind of absurd when we do that.[4] *What* idiomatic and highly standardized metaphors genuinely assert can be paraphrased without residue – although not with an eye to a *common property* that the various objects to which the metaphor applies have (as Yablo and Toon suggest), but to the *context* in which the metaphor is used.

Take the following example (from Grant, 2013, pp. 126–131). It makes a difference *who* calls someone a gorilla and under what *circumstances*. When Muhammad Ali called Joe Frazier a "gorilla", that was an insult; but coming from someone well-versed in primatology, it is a compliment. The point is that "*x* is a gorilla" can be paraphrased, depending on the *context*, as *x* is "enormous", "hairy", "powerful", "brutal", or "gentile" and "peace-loving". We do not use this metaphor "to represent people as having a property that cannot be attributed to them by speaking literally" (Grant, 2013, p. 128). It is not that the metaphor applies to objects because they have a common property; but that the same metaphorical expression can be apt due to different properties, which can also be described in literal terms.

Dead, idiomatic, and standardized metaphors would allow for Toon to claim that FP explanations are GCEs, since in these cases the explanation can refer to real events and real causal connections between them. But they can do this precisely because they are not representationally essential: they are either literal expressions like category i) or can be paraphrased in non-metaphorical ways like category ii).

This leaves us with category iii) i.e., poetic metaphors as the only remaining candidates for the role of REMs: they cannot be paraphrased without losing their function. In the case of poetic metaphors, every attempt at paraphrasing leaves out something that is essential to these metaphors.[5] And this is what makes them essential in the given context: no other construction can convey what they communicate. No paraphrase can save the framing effect, the insight (if any), the affective surplus etc. of metaphors like "Light is but the shadow of God" (Thomas Browne) or

> For most of us, there is only the unattended
> Moment, the moment in and out of time,
> The distraction fit, lost in a shaft of sunlight,
> The wild thyme unseen, or the winter lightning
> Or the waterfall, or music heard so deeply
> That it is not heard at all, but you are the music
> While the music lasts. (T.S. Eliot)

But poetic metaphors pose at least the following three problems for Toon. First, poetic metaphors are instantly recognizable by competent speakers. Poetic metaphors, unlike dead and idiomatic ones, cannot succeed if they are unnoticed. If they cannot succeed when unnoticed, they cannot play the part of Toon's FP metaphors. And when they succeed then, to speak Toon's language, we do ordinarily feel that we are speaking, or being spoken to, metaphorically. Consequently, either Toon's FP metaphors are unnoticed, but then they are dead or idiomatic and so not REMs, or they are REMs (poetic) but then must be noticed in order to succeed. Either REMs or Toon's response to the phenomenological objection must go.

Secondly, poetic metaphors cannot serve as a model for the kind of metaphor that Toon needs in his account. Poetic metaphors do not have Waltonian *principles of generation*. In the sense as Walton (1990, pp. 40–41) defines, and Toon (2023, p. 23, p. 48) invokes them, these principles are "rules" or "conditional prescriptions about what is to be imagined in what circumstances". They are thus systematic connections between actual circumstances in the real world and fictional truths. Or to put it differently, a principle of generation makes it possible to derive fictional truths from truths in the real world. When children pretend that the baby doll is sleeping because its eyes are closed, they follow one such simple rule (pp. 15–16). Poetic metaphors do not have such principles. This is the reason why poetic metaphors strike us as artistically creative, can be strikingly effective, but cannot be produced systematically and cannot be paraphrased. These features reflect the lack of principles of generation for poetic metaphors. And this is the reason why Walton's analysis of these principles cannot be applied to them. Not surprisingly, because Walton's analysis is about principles for generating fictions and not poetic metaphors.[6]

Now Toon's FP metaphors *have* principles of generation, and consequently they lack all these features of poetic metaphors. For him these principles are "complex, nuanced, and largely tacit rules that govern our folk psychological practices". How things are with the "props" (recall: people and their behavior) plus these principles "generate the content of our make-believe" (p. 23). This content "possesses a certain kind of 'objectivity'" (p. 16).[7] This objectivity is inconceivable if the principles *qua* rules are not common knowledge to the pretenders, because then they could not come up with sustainably coherent moves in the game of make-believe.

For the present readers, it is unclear why, if we have such principles, it is impossible to paraphrase FP ascriptions into behavioral or some other descriptions – just like in the case of the baby doll. Toon's principles are systematic to the extent that they ground our "reasonably coherent set of rule-governed practices for attributing mental states to people based on their behavior" (p. 48). We can easily agree that these rules "are complex and difficult to formulate explicitly" (p. 48), but explicit formulation must not be impossible because they are there: they must be internally represented in order for a competent speaker to be able to play the game of make-believe. And as such, it must be possible to make them explicit with some theoretical effort – just like the internally represented rules of grammar can be made explicit by linguists.

It seems Toon cannot have REMs *plus* principles of generation. If we have metaphors with such principles, then they are not representationally essential, but either dead and as such literal expressions, or standardized and idiomatic and as such paraphrasable. Those metaphors

that have no such principles are not Toon's FP metaphors: the latter do have principles of generation, so they cannot be representationally essential.

Thirdly, if FP explanations are GCEs, then their relata must stand in the relation specified in the explanations. This is a minimal criterion of extensional adequacy that any (and not only genuine and not only causal) explanation must meet. Apparently, Toon's FP is conformable with some version of this criterion. He thinks we make "genuine assertions" that are made true or false by "a complex set of facts" about behavior (p. 23), and these assertions "help us to pick out genuine facts about people's behavior" (p. 81). If FP is indeed like this, there is nothing to worry about its causal character.

But if FP tries to achieve this by REMs, then it is doomed because REMs are unsuited to pick out genuine facts. Notice that we advocate loose standards here, and we could accept the principles of generation for ensuring extensional adequacy. But REMs (*qua* poetic metaphors), as we have argued, do not conform to principles of generation, and we cannot see how they could in any other way be systematically connected to "genuine facts about people's behavior" in Toon's framework. And vice versa, if a metaphor is *systematically* connected to genuine facts, then it is not a REM because then it can be paraphrased in terms of the facts themselves. This again shows that if you are committed that a discourse relies on REMs, then you cannot consistently insist on the possibility of GCEs in the same discourse.

5. But is it fictionalism?

Up to this point, we have argued that REMs and GCEs are incompatible. But why would anyone expect GCEs from a *fictional* discourse? Or alternatively, why would anyone who wants to preserve GCEs in any field of discourse want to advance a *fictionalist* interpretation of that discourse?

Inspired by Gideon Rosen (2005, p. 14) let us suggest three sensible commitments that a fictionalist about a given discourse should take in order for his position to be discernible from alternatives.[8] Accordingly, the claims of the fictionally interpreted discourse

(i) are genuine representations of how things stand, and so are truth-evaluable;
(ii) are to be interpreted non-reductively at face value;
(iii) do not have epistemic aims, but they have other, non-epistemic virtues.

The third is the distinctive commitment, but the three in conjunction uniquely identify the fictionalist stance.

In Toon's account, FP aims at GCE. Any account of a discourse that takes its aim to be GCE, we submit, fails to qualify as fictionalist. GCE provides causal information about the explanandum. Causal information is possible only if extensional adequacy is ensured in some way: causes must be picked out and stand in the specified relation to their denoted effects. A discourse aiming at GCE must aim at explanations that track the relation of causes to effects, so it must be in the business of telling truths.

Almost everyone thinks that FP is in this business, so almost no one qualifies as a mental fictionalist. Apart from Toon, of course, the most plausible candidate is Daniel Dennett (e.g., Dennett, 1987, 1991), and Toon speaks a remarkably similar language to Dennett at times. He even claims that his fictionalism "can make sense of aspects of Dennett's view that his critics have found most problematic" (p. 41). Toon's making sense, however, reveals why his account fails to qualify as fictionalist – just as Dennett's account fails to do so.[9]

The heart of the matter is this.[10] If FP explanations are GCEs, then their cardinal virtue is epistemic, and there is no need for their distinctively fictionalist account. Even if they are "not straightforwardly true" (p. 5), because that is too much to be hoped for, they are at least *systematically connected* to truths about behavior. But "fictional truths" so generated are not genuinely fictional but are just truths generated from other truths (via e.g., principles of generation). As such, they are useful *instruments* for the purposes of explanation and prediction with ostensible epistemic virtues and connections to other truths.

Toon's (p. 44) discussion of Dennett's centers of gravity and patterns of behavior make this clearer. Centers of gravity are not metaphors. Instead, they are mathematical constructs calculated according to explicit "principles of generation": there is a literally true description of geometrical properties of an object, and there are principles as to how to calculate centers of gravity from those truths. No metaphor here, but a systematic connection of literal truths that explains the instrumental, explanatory, and predictive benefits of such constructs.

The same holds for Dennett's patterns of behavior with a distinctive FP flavor. These patterns are real, but they are discernible only from the intentional stance. The intentional stance is a specific way of looking at pieces of behavior with an eye to their coherence and rationality. The stance gives us the principles for connecting the dots and revealing a pattern, but the principles are lax, vague, more complex than those revealing centers of gravity, and can be deployed idiosyncratically. Therefore, the dots can be connected in several incompatible ways resulting in conflicting interpretations of a person's behavior.

Toon compares his metaphor of "angry clouds" to Dennett's patterns (pp. 44–45). Rightly so, but this reveals that there are systematic principles for deriving the metaphor from observable patterns, and the patterns can be recognized only against the background of the rules of "this particular game" – the angry clouds game, that is. The rules here, as in the case of Dennett's intentional stance, are implicit, but they can, at least in principle, be made explicit; they systematically connect certain literal truths (features of my visual perception) to the metaphor (my judgment that these clouds are angry) and make it true. There seems to be nothing mysterious here if compared to my perception of a giraffe: I can no more tell the clues from which I judge that "a giraffe is in the veldt", than the clues from which I judge "the clouds are angry". It is for these systematic connections that we can usefully invoke the metaphor in GCEs. No need for fictionalism here, no need for pretense, and not much work left for metaphor.

Toon can still argue that for the products of this process we use concepts adopted from, or at least inspired by, some other domain, and in this sense they are metaphorical. But these are not REMs, because they can, in principle but unproductively, paraphrased in terms of literal truths about behavior and "principles of generation". Alternatively, they could be replaced by other metaphors, because a metaphor so understood is just a shorthand for structural features of complex behavioral facts or literal truths imported from another domain (public representations, in Toon's case, but it could be, say, the domain of computational processes or something else). These metaphors are only convenient tools for higher-order talk about the products of this inferential process for the purposes of explanation or prediction.

A distinctively fictionalist account of FP in the spirit of the three criteria proposed above is motivated only if one wishes to *deny* that FP has epistemic virtues, but this cannot be the case if FP aims at GCEs. The only alternative form of mental fictionalism that Toon (2023, pp. 46–48) engages with in the book, prefix fictionalism, is not of this kind either.[11] This version also agrees that FP is a predictively and explanatorily oriented discourse (see Wallace, 2022, pp. 35–36), that presupposes systematic connections between literal truths and "fictional truths". Otherwise, success would be an unexplained cosmic coincidence. If FP explanations are GCEs, then it seems very problematic to separate predictive and explanatory success, as Toon (2023, pp. 99–103) seems to be inclined. If we can pick out the genuine causes of behavior in our FP explanations, it cries for an explanation if its predictive success is merely "limited" (p. 103). If those causes are genuine, even if they are picked out metaphorically, why cannot we invoke them with more predictive success?

An alternative to both prefix- and pretense-fictionalism that Toon knows but surprisingly does not discuss is affective fictionalism.[12] This proposal

suggests that the proper domain of FP is not the explanation and prediction of behavior but its evaluation and interpersonal orientation in the social world.

The affective fictionalist draws attention to the fact that we are typically interested in FP explanations where a simple "because" "Mark believes that the No. 73 goes to Oxford Street" or "Ruth wants a cup of coffee" (p. 28) will not do. Everyday cases of social interaction do not stimulate our FP sensitivity. More often than not, we just rely on routines, frames, and subpersonal processing. These cases can be easily construed for philosophical purposes as involving complex FP attributions, but we do not need to overintellectualize them that way.

Outside philosophy, we typically deploy FP concepts when we need to reflect on why people have done something, or have done it in a specific way, or what they will do.[13] When we ascribe mental states in such cases, we come up with an interpretation of someone's behavior that resolves some tension arising from a lack of understanding. Thus, we make the behavior relatable or antipathic by the interpretations with which we portray its seemingly causal background.

So construed, FP interpretations express our sentiments that arise in these problematic social situations and convey them in accordance with the conventions of FP interpretations. These conventions are subject to change over time and can differ from community to community, but while in place, they stabilize the meaning of FP terms and the sentiments that their various constructions can express. In this respect, FP resembles to the conventions of musical composition in happier times (see Demeter, 2013).

These interpretations are not systematically connected to facts about the behavior of the interpreted person but are systematically connected to the interpreters' sensitivity, conceptual resources, and their previously given interpretations. How those interpretations are received is connected to the same features of the listeners. FP so conceived is a tool of conveying and tuning socially significant sentiments. It fulfills a social function and has a motivational role: if an FP interpretation is convincing, it tunes interpretive sensitivities and the sentiments of the listeners toward the object of the interpretation. Thus, FP discourse does not aim at epistemic benefits but at success in the social world: at getting along well, forming alliances, and orienting others to see some parts of the social world in the way we want them to.

Affective fictionalism also draws on a theory of metaphors, but *without* claiming that FP terms are metaphors. Here, FP interpretations behave *like* Davidson's (1978) metaphors (irrespective of whether this is a good theory of metaphors or not) in that they have semantically evaluable content (they genuinely represent people as having mental states), but

what matters is that they carry a non-semantic, affective content, and their social virtues are due to *this* content. Consequently, FP interpretations do not aspire to be GCEs. In some respects, they are *like* poetic metaphors: they cannot be effectively paraphrased and have no replacement in our life, so they are like REMs. But they do not face the dilemma that we have explored above and that Toon cannot avoid because they do not aim to give GCEs.

Moreover, with Davidson's help, affective fictionalism avoids the *prima facie* most implausible feature that Toon's account inherits due to its insistence on Walton's theory, namely that in ascribing mental states to others we *pretend as if* they had them, but without knowing that we are pretending. The affective fictionalist does not think we pretend; we just do not know what we are actually doing in FP discourse when we do what we *prima facie* think we do. It seems very hard to maintain community-wide practices of pretense without anyone but Toon realizing that we are pretending. Maybe this is because it is very hard to pretend without knowing what it would be like not to pretend. It seems almost as hard as to lie when one does not only not know but cannot even conceive the truth.

6. Conclusion

Toon thinks that FP is epistemically indispensable: there is no more adequate way to understand and explain someone's behavior than relying on FP. He also thinks that FP plays this role by providing GCEs in terms of REMs. In this paper, we have advanced three challenges against this view. First, we could see no way to literally misconstrue the alleged metaphors of FP, i.e., FP could not pass the test for nonliteralness. Second, it seems that only with an undifferentiated concept of metaphor can one pretend to reconcile REMs and GCEs. If seen through the lens of a slightly more differentiated concept, then it turns out that there is no way of unpacking the idea of REM so as to be compatible with Toon's commitment of unparaphasability and GCE. And finally, we argued that a fictionalist account cannot preserve the alleged epistemic virtues of FP without collapsing into some form of instrumentalism. Mental fictionalism worthy of its name requires one to resign to the loss of FP's alleged epistemic benefits.

Notes

1. With respect to universities, of course we are hinting at Ryle (1949/2009, p. 6).
2. On p. 29 Toon seems to commit this fallacy by implying that it is due to metaphor that we can ask silly questions: "To ask where beliefs *are* [. . .] is to ask a silly question – one that misunderstands the metaphorical nature of our talk about mental states".

3. This can happen by accident, in ways similar to Gettier cases. Non-accidentally and non-systematically, it is of course possible to give an explanation based on make-believe that by accident proves to be a GCE.
4. Toon (2023, p. 26) reads Ryle's critique of Cartesianism as being based on this insight. However, Toon differs from Ryle in concluding that this category mistake is not due to a mistaken theory, but rather a natural consequence of the metaphorical nature of the discourse.
5. At this stage it doesn't matter much what this "something" is. What matters is that it is something essential for the metaphor to fulfil its function in the given context. Here we are alluding to Davidson (1978).
6. This also reflects the fact that metaphor and fiction do not go hand in hand. Metaphors are not necessary for fictional contexts, and fictional contexts can be generated without metaphor.
7. In this sentence objectivity is in scare quotes. By the next paragraph Toon's confidence has grown and the scare quotes disappear.
8. These commitments are inspired by, but not identical with, Rosen's. The third commitment here is more radical than Rosen's third. His fictionalist only denies that the fictionally interpreted discourse aims at *truth* and he is willing to grant other epistemic virtues to the discourse in question. On Rosen's criteria, van Fraassen's constructive empiricism (or Dennett's instrumentalism) can count as a fictionalist stance; on the criteria suggested here they do not qualify as fictionalist positions.
9. For Dennett's (2022) hesitation
10. See Demeter (2022a) for a more detailed exposition.
11. For prefix-fictionalism, see Parent (2017, pp. 226–230) and Wallace (2022).
12. See Demeter (2009, 2013, 2022a, 2022b). For a parallel discussion of pretense and affective mental fictionalism see Bourne and Caddick Bourne (2020).
13. For a very useful detailed discussion of broad and narrow construals of FP see Bermúdez (2005), Chapter 7, esp. pp. 176–178).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the MTA-BTK Lendület Value-Polatizations in Science Research Group and NKFI K 146597.

ORCID

László Kocsis http://orcid.org/0000-0003-3147-2281
Krisztián Pete http://orcid.org/0000-0002-8454-195X

References

Bermúdez, J. L. (2005). *Philosophy of psychology: A contemporary introduction*. Routledge.

Bourne, C., and Caddick Bourne, E. (2020). Folk stories: What has fiction to do with mental fictionalism? In B. Armour-Garb and F. Kroon (Eds.), *Fictionalism in philosophy* (pp. 168–186). Oxford University Press. https://doi.org/10.1093/oso/9780190689605.003.0009

Burgess, J. P. (2004). Mathematics and bleak house. *Philosophia Mathematica*, *12*(1), 18–36. https://doi.org/10.1093/philmat/12.1.18

Davidson, D. (1978). What metaphors mean. *Critical Inquiry*, *5*(1), 31–47. https://doi.org/10.1086/447971

Demeter, T. (2009). Folk psychology is not a metarepresentational device. *European Journal of Analytic Philosophy*, *5*(2), 19–38.

Demeter, T. (2013). Mental fictionalism: The very idea. *The Monist*, *96*(4), 483–504. https://doi.org/10.5840/monist201396422

Demeter, T. (2022a). Psychological understanding. In A. T. Tuboly (Ed.), *The history of understanding in analytic philosophy: Around logical empiricism* (pp. 229–259). Bloomsbury Academic. https://doi.org/10.5040/9781350159235.ch-010

Demeter, T. (2022b). A mental fictionalism worthy of its name: A plea for affective storyism. In T. Demeter, T. Parent, & A. Toon (Eds.), *Mental fictionalism: Philosophical explorations* (pp. 86–103). Routledge. https://doi.org/10.4324/9781003091073-6

Dennett, D. (1987). *The intentional stance*. MIT Press.

Dennett, D. (1991). Real patterns. *The Journal of Philosophy*, *88*(1), 27–51. https://doi.org/10.2307/2027085

Dennett, D. (2022). Am I a fictionalist? In T. Demeter, T. Parent, & A. Toon (Eds.), *Mental fictionalism: Philosophical explorations* (pp. 352–362). Routledge. https://doi.org/10.4324/9781003091073-23

Grant, J. (2013). *The critical imagination*. Oxford University Press.

Parent, T. (2017). *Self-reflection for the opaque mind. An essay in neo-sellarsian philosophy*. Routledge.

Rosen, G. (2005). Problems in the history of fictionalism. In M. E. Kalderon (Ed.), *Fictionalism in metaphysics* (pp. 14–64). Clarendon Press. https://doi.org/10.1093/oso/9780199282180.003.0002

Rosen, G., and Burgess, J. P. (2005). Nominalism reconsidered. In S. Shapiro (Ed.), *The Oxford handbook of philosophy of mathematics and logic* (pp. 515–535). Oxford University Press. https://doi.org/10.1093/0195148770.003.0016

Ryle, G. (1949/2009). *The concept of mind*. Routledge.

Toon, A. (2023). *Mind as metaphor*. Oxford University Press.

Wallace, M. (2022). Mental fictionalism. In T. Demeter, T. Parent, & A. Toon (Eds.), *Mental fictionalism: Philosophical explorations* (pp. 27–51). Routledge. https://doi.org/10.4324/9781003091073-3

Walton, K. (1990). *Mimesis as make-believe: On the foundations of the representational arts*. Harvard University Press.

Walton, K. (1993). Metaphor and prop oriented make-believe. *European Journal of Philosophy*, *1*(1), 39–57. https://doi.org/10.1111/j.1468-0378.1993.tb00023.x

Yablo, S. (1998). Does ontology rest on a mistake? *Proceedings of the Aristotelian Society, Supplementary Volume*, *72*(1), 229–261. https://doi.org/10.1111/1467-8349.00044