9

Higher-Order Metaphysics and Propositional Attitudes

Harvey Lederman

1. Introduction

The Milesian monists held that there is only one kind of thing: Thales said it was water, Anaximenes air, and Anaximander the "unlimited". Later thinkers came to reject or at least refine this bold idea. Atomists like Democritus proposed that, while everything is made from atoms, the atoms come in different, irreducible, inconvertible kinds. And, in their different ways, Plato and Aristotle thought that there must be forms in addition to material stuff.

But arguably all of these ancient thinkers would have agreed with a more abstract form of monism, which has dominated analytic metaphysics since Quine. This position is not the jaw-dropping if splendid idea that everything is water but a more abstract one, the claim that everything there is—electrons, chairs, nations, properties, whatever it might be—is a single kind of entity: an object, a thing. For a time, this position—which we might call "objectual monism"—was so dominant that it was rarely defended, and was simply assumed. But well before Quine, an alternative had already been proposed. In 1891, articulating a position he seems to have held for some time, Frege wrote that there is a distinction "founded deep in the natures of things" between first-level and second-level concepts ("Function and Concept", Frege et al., 1970, p. 41), and he quite clearly held the same for the distinction between objects and concepts (or, as I will say, "properties"). Putting it very roughly—and asking, like Frege himself, to be interpreted with a "pinch of salt" ("On Concept and Object", Frege et al. 1970, p. 54)—he proposed that the atoms of reality come in different, irreducible, inconvertible kinds.

Frege often described his idea by contrasting "complete" entities—objects—with "unsaturated" ones—properties. But in the catacombs where Frege's idea lived while Quine dominated the world out of doors, this contrast was no longer seen as the heart of the approach. Instead, the emphasis fell on a different aspect of Frege's thought, that syntactic categories in language correspond, very roughly at least, to ontological ones (Prior, 1971). Just as names, predicates, and sentences play different roles in language, so too (the rough idea goes) entities in the

Harvey Lederman, *Higher-Order Metaphysics and Propositional Attitudes* In: *Higher-Order Metaphysics*.
Edited by: Peter Fritz and Nicholas K. Jones, Oxford University Press. © Harvey Lederman 2024.
DOI: 10.1093/oso/9780192894885.003.0009

corresponding ontological categories—objects, properties, propositions—play different roles in the world.

This style of view—which I'll call "neo-Fregean pluralism", or sometimes, very loosely "higher-order metaphysics"—has recently surfaced from its underground hideaway to become the basis of a full-blown evangelical movement.[1] The movement's appeal derives in part from dissatisfaction with monism's most basic claims. The monist says that if "is happy" or "The sky is blue" have meanings, their meanings must be *things* like electrons, chairs, numbers, or sets. But on first encounter with the theory of meaning, this claim is apt to seem bizarre. Could "is happy" really mean a *thing*? Similarly, the monist says that if *is happy* or *the sky is blue* are at all, they must be objects, things. But on first encounter with the theory of properties or propositions, this claim is apt to seem bizarre. Could *is happy* or *the sky is blue* really be a *thing*? Some of us lived so long with the dogma of objectual monism that we forgot how strange these doctrines seemed at the start. But once remembered their oddity is hard to ignore.

The apostles of higher-order metaphysics have won converts by stirring these nearly dead embers of doubt to life. They say that there can in some sense be meanings of predicates and sentences, even if those meanings are not to be found in the realm of things. According to them, these meanings are (speaking roughly) *sui generis* entities, members of ontological categories distinct from the category of things. They are in some ways similar to the monist's properties and propositions. But unlike properties or propositions as imagined by the monist, these *sui generis* entities are not things of any kind, whether abstract or concrete.

With their audience on fire from this impassioned appeal, the evangelists shift to a calmer, more theoretical key. These first intuitive ideas, they say, are not naive, confused thoughts of which no sober sense can be made. The outlook can be stated exactly in the elegant, well-understood framework of the simple theory of types. It yields a strong, attractive theory of what *sui generis* properties there are. And it dissolves, or at least makes progress on dissolving, gnarly old quandaries—whether about third men, relational regresses, or even where properties could be—which have beset objectual monism essentially right from the start (see e.g. Jones, 2018).[2]

[1] Caplan (2011) calls a related view "Fregean ontological pluralism"; Trueman (2021) uses "Fregean realism". As I understand it, this view is a species of the genus of ontological pluralist views, described in (e.g.) Turner (2010), McDaniel (2017). But discussions of pluralism in general typically focus on a quite different species of this genus, according to which (as I would say) everything there is is an object, although these objects enjoy different ways of being. It's beyond the scope of the present paper to say why one paraphrase rather than another might be appropriate (see Caplan, 2011 for some discussion), but I'll focus on the neo-Fregean form here.

[2] More concretely: some monists hold that, if there are properties like redness, there must be a special "glue" of instantiation which can bind them to ordinary things like apples or oranges but which cannot bind apples to oranges. Since monists hold that redness, apples, and oranges are all things of (in some sense) the same sort, they must explain why the glue sticks to redness but not to apples. Pluralists, by contrast, have a built-in account: for them, entities in the category of *sui generis* properties combine (as it were, by nature) with objects in something like the way that predicates combine with names, while entities in the category of things or objects cannot combine with one another in the same way. Similarly, monists who hold that properties are intrinsically "sticky" must explain in what way they differ from other objects, while pluralists have a ready made account.

The excitement about this style of view has been palpable in the last few years, and in my view, justly so. But sometimes its proponents can make it seem as though the right kind of pluralism can dissolve all the old problems with properties and propositions. I wish that this were true, but here I'll argue that, at least in one important case, it's not.

According to a prominent form of pluralist, there is (roughly speaking) a *sui generis* category of entities denoted by sentences that play some of the roles propositions play for monists. It is natural for those who endorse this view to hope that they can appeal to these distinctive entities to give a simple account of propositional attitudes like belief. On this view, for Alice to believe that some rabbits can speak is (roughly) for Alice to stand in a relation to the *sui generis* entity denoted by "some rabbits can speak". In fact, Arthur Prior, one of the most prominent neo-Fregean pluralists, seems to have taken the availability of this straightforward account of propositional attitudes to be a key attraction of his position (Prior, 1971, cf. Prior, 1963).

But here I'll argue that this straightforward account cannot be straightforwardly accepted. I'll start by showing how the view conflicts with core principles of higher-order logic, given natural judgments about propositional attitudes. This observation will already undercut an important motivation for higher-order metaphysics. But I'll argue that there are deeper challenges ahead: that the attitudes may erode the foundations of this pluralist approach.

Section 2 introduces neo-Fregean pluralism and describes how it promises a straightforward account of propositional attitudes. Section 3 shows how, given natural assumptions about propositional attitudes, this straightforward view conflicts with core principles of higher-order logic. Section 4 then describes the main challenges of the paper. (An impatient reader who wants to get a quick sense for the main ideas may wish to skip ahead and read this section first.) Section 5 further develops the challenge, arguing that higher-order metaphysicians shouldn't set the attitudes aside as entirely irrelevant to more metaphysical concerns. Section 6 considers whether higher-order metaphysicians should respond to the main challenges by embracing a radical theory of propositional attitudes. Section 7 concludes with general remarks on the relationship between the main themes of the paper and questions about "logical omniscience" and "hyperintensionality". An Appendix considers how some "opacitist" views fare with respect to the central challenges.

2. Naive Relationism and Higher-Order Metaphysics

This section motivates and describes a position about propositional attitudes which I call "naive higher-order relationism" (or "naive relationism" for short). On the way to doing so, I introduce a higher-order language and discuss why languages of this kind are so important for neo-Fregean pluralists. Readers already familiar with higher-order metaphysics may wish to skip past this material to the penultimate paragraph of the section, where naive relationism is introduced.

In the sentences "Alice hopes that some rabbits can speak", "Alice fears that some rabbits can speak", "Alice knows that some rabbits can speak", and "Alice believes that some rabbits can speak", the verbs "hope", "fear", "know", and "believe" are all followed by the word "that" (a "complementizer") and then a sentence ("some rabbits can speak"). When these verbs (and verbs like them) are used with such "sentential complements", they are called "propositional attitude verbs" and they express (or express something closely related to) underlying states of mind, which are themselves known as *propositional attitudes*. It is controversial how to characterize the class of propositional attitudes exactly, but for my purposes below I won't need to; it will be enough to appeal to canonical examples like these. For simplicity, in fact, I'll mostly focus on the one example of belief.

In the sentence "Alice loves Charlotte", the word "loves" expresses a *relation* between Alice and Charlotte. Similarly, in the sentence "Alice believes that some rabbits can speak", "believes" seems to express a relation, in this case between Alice and whatever is denoted by "some rabbits can speak". *Relationists* about propositional attitudes take inspiration from this apparent similarity between loving and believing. They hold that propositional attitudes are relations between attitude-holders and something else, which we might call the "objects" of the attitude, or, picturesquely, an "object of thought".

Relationism is an extremely attractive position, and arguably the most popular view today about the metaphysics of propositional attitudes. But relationists must answer a hard and important question: what are these "objects of thought"? The "objects" have some curious properties. Most notably, if Alice believes that some rabbits can speak, plausibly what she believes has a salient property in the situation in which we find ourselves (where no rabbits can speak) that it would not have had in a different situation (if some rabbits had been able to speak). It is natural to describe this difference in English by saying that what Alice believes is in fact incorrect, inaccurate, or false, but that it would have been correct, accurate, or true in an alternate scenario. Below, I'll refer to this contrast using the terms "true" and "false". A satisfactory account of the objects of thought must explain how they can be true or false in this sense. At the end of the section, we will see how higher-order metaphysicians might use their distinctive ontology to give such an account. In particular, we'll see how they might do so if they held (very roughly) that the "objects of thought" are the *sui generis* entities which higher-order metaphysicians take to be denoted by sentences in their higher-order language. (For this view, see e.g. (arguably) Ramsey, 1931, pp. 142ff.; Prior, 1963; Prior, 1971; Rayo and Yablo, 2001; Rosefeldt, 2008; Trueman, 2021, Ch. 12.4; Jones, 2019, §6; D'Ambrosio, 2023.)

Higher-order metaphysicians use a special kind of "higher-order" language to articulate their metaphysical views exactly. So, to state the position about propositional attitudes I've just sketched, we have to learn a little about these languages. Here I will use a higher-order language in which every expression has a unique "syntactic type". The "types" will be the smallest set containing e, t, and, for all types σ and τ, a further type $\sigma \to \tau$. These types encode which expressions can grammatically combine with which others. In particular, expressions of type $\sigma \to \tau$ combine with expressions of type σ (their "arguments"), to produce expressions of type τ, and this is in fact the only way that expressions can combine with others to create new expressions. We assume (at least for now) that English names like "Alice" are translated into the higher-order language as expressions of type e and that sentences like "Alice is happy" are translated as expressions of type t. Predicates like "is happy" are then translated as expressions of type $e \to t$, so that, given our rule, the translation of "is happy" ($H^{e \to t}$, with the superscript indicating its type) can combine with the translation of the name "Alice" (a^e) to produce the translation of "Alice is happy" (Ha, an expression of type t). (In general I'll drop the type-superscripts after they first appear.) Similarly, unary sentential operators like "it's not the case that" will be translated as expressions of type $t \to t$, so that they can combine with sentences to yield other sentences. For instance the translation of "it's not the case that" ($\neg^{t \to t}$) can combine with the translation of "Alice is happy" (Ha) to produce another sentence, the translation of "it's not the case that Alice is happy" ($\neg Ha$, an expression of type t). Finally, since expressions can only combine in line with the rules associated with their type, the translation of "it's not the case that" ($\neg$, $t \to t$) cannot combine with the translation of "Alice" (a, e) to produce a grammatical expression, and similarly the translation of "is happy" (H, $e \to t$) cannot grammatically combine with the translation "Alice is happy" (Ha, t) and so on. (As these examples illustrate, English word-order does not always correspond to how I'll regiment argument structure in higher-order translations.)

Because I find it easier to work with, I have chosen a higher-order language with the artificial feature that each expression can only take one argument. (This is reflected in the fact that all types other than e and t have the form $\sigma \to \tau$, with only a single type on the left of the arrow.) But we can still translate expressions like "loves" or "and", which intuitively take two arguments, into expressions of our language, which take their arguments one at a time. We simply have to pick an order for their relevant arguments. For instance, in the case of "loves", we can use a term with type $e \to (e \to t)$, $L^{e \to (e \to t)}$, which combines first with an expression of type e (e.g. a, our translation of "Alice") to produce an expression of type $e \to t$ (La, "Alice loves"), which can in turn combine with an expression of type e (e.g. c, translating "Charlotte") to produce an expression of type t ($(La)c$ "Alice loves Charlotte"). (After this I'll often drop parentheses.) Similarly, "and" will be translated by a term with type $t \to (t \to t)$, $\wedge^{t \to (t \to t)}$, which first combines with a sentence (e.g. Lac for "Alice loves Charlotte") to produce an expression with the type of a unary sentential operator ($\wedge(Lac)$ for "Alice loves Charlotte and", of type $t \to t$), which in turn combines with another sentence (e.g. Ha for "Alice is happy") to produce a sentence $((\wedge(Lac))Ha$, which I'll typically abbreviate using the standard infix notation as $Lac \wedge Ha$, for "Alice loves Charlotte and Alice is happy"). Although, as I've said, I find this "functionally typed" language (where expressions take one argument at a time) easier to work with, the fact that it forces us to make such arbitrary choices does mean that it seems less metaphysically perspicuous than a "relationally typed" language in which expressions can take multiple arguments at a time. (Could there really be a fact about whether love takes the lover or the beloved first?) Fortunately, for the most part, everything we do in a functionally typed language can be seamlessly translated into a more perspicuous relational one (see e.g. Dorr, 2016, Appendix). So we can use the more convenient language mostly without harm, and I'll continue to do that here.

On its own, the fact that someone speaks a higher-order language tells us little about their metaphysical views. But here I will be interested instead in someone who takes the syntax of this language to correspond in an important way to reality. Our higher-order metaphysician holds (roughly speaking) that, for every syntactic type, there is a corresponding ontological category. They hold (still speaking roughly) that expressions of type t (which I'll call "t-propositions") denote entities belonging to a different ontological category than expressions of type $e \to t$ (which I'll call "et-properties"), and that both in turn are entities in different categories than those denoted by expressions of type e ("objects", "things").

These paraphrases, which use the expression "ontological category", capture something important about neo-Fregean pluralism. But they are, strictly speaking, incorrect. It will be worth pausing for a moment to see why, since doing so will help to clarify the special importance of higher-order languages for neo-Fregean pluralists, and this will play a key role in the main challenges I'll come to in Section 4.

Consider, as a warm-up, the expression "is an entity". This expression is most naturally translated into our higher-order language by an expression with syntactic type $e \to t$ since, like "is happy", it combines with a name (e, e.g. "Alice") to yield a sentence (t, "Alice is happy", "Alice is an entity"). But given this translation, straightforward translations of the sorts of things I've been saying won't express what the higher-orderist believes. Suppose I tried to articulate the view by writing "being happy is an entity of an altogether different kind than a flamingo". This sentence will be grammatical, on the natural supposition that "being happy" is an expression of type e, but it will not accurately express the neo-Fregean pluralist's view. According to them, the interesting entity in the vicinity of being happy is not an object, denoted by an expression of type e, but an entity which can only be expressed by expressions of higher type, most obviously by the ($e \to t$) translation of "is happy". This problem might seem fiddly, and one might think we could get around it by saying instead "is happy is an entity of an altogether different kind than a flamingo". But given our rendition of "is an entity", this sentence is also ungrammatical: expressions of type $e \to t$ can only take expressions of type e as arguments; they cannot grammatically combine with other expressions of type $e \to t$. (In the introduction, I used an italicized "*is happy*" as a half-way house nominalization, in a vain attempt to make this read like English, but that was a fudge.) These points don't just hold for "is an entity", they hold for key expressions like "is a proposition", "is a property", and, crucially, for "is an ontological category" and "belongs to an ontological category" as well. Each of these expressions will be translated into the higher-order language as an expression of type $e \rightarrow t$, and so can't be used to say what the higher-orderist wants to say. In particular, sentences like my "neo-Fregeans hold that there are different ontological categories" can't be exactly right.

What will the neo-Fregean say instead of these rough claims? They can articulate their position exactly using the higher-order language itself, and in particular using two last pieces of special vocabulary: higher-order identity and higher-order quantification. In addition to the identity-symbol of first-order logic, which will be translated into our higher-order language as "$=_e$", an expression of type $e \rightarrow (e \rightarrow t)$ (like the translation of "loves"), our higher-order language will also have, for every type σ, an expression $=_\sigma$, of type $\sigma \rightarrow (\sigma \rightarrow t)$, allowing us to make grammatical identity statements featuring higher-type expressions. For instance, we can use these expressions to form sentences like (the true) "$H =_{e \rightarrow t} H$" (roughly, "to be happy is to be happy") or (the false) "$Lac =_t Lca$" (roughly "for it to be the case that Alice loves Charlotte just is for it to be the case that Charlotte loves Alice"). Like identity, the notion of quantification can also be generalized to types other than e. In addition to a quantifier associated with type e, which, like the quantifier of first-order logic, allows us to generalize in the syntactic position of names, we will have, for every syntactic type σ, a quantifier $\forall_\sigma$, which allows us to generalize in the syntactic position of expressions of type σ. For instance, we can now form sentences like "$\forall_t p^t \neg p$" (roughly, "everything is not the case") and "$\forall_{e \rightarrow t} F^{(e \rightarrow t)} Fa \rightarrow \neg Fc$" (roughly, "every way Alice is, Charlotte is not"). (As you can see, variables in the language will have types too, just like the constants I focused on above.)[3]

We're now at last in a position to state the promised precise claims. In first-order logic, it's standard to express existence using identity and quantification, and we can translate those claims into our higher-order language as well. For instance, $\exists_e x x^e =_e a$ can translate "Alice exists". In the higher-order language we can also make such claims using expressions of higher types, for instance, $\exists_{e \rightarrow t} F^{e \rightarrow t} F =_{e \rightarrow t} H$ (roughly, "is happy exists"), $\exists_t p^t p =_t Ha$ (roughly, "Alice is happy exists"). (From now on, as here, I'll put in enough type-markers to make sure it's clear how to read the relevant sentences/schemas, but I won't follow strict

[3] Officially, the quantifiers for type σ, $\forall_\sigma$, are treated as constants of type $(\sigma \rightarrow t) \rightarrow t$, and λ-abstracts (which I'll introduce below) are used to handle variable binding. To say "everything is happy" we'd officially write $\forall_e H$ (think of it as "is happy is e-universally instantiated"). Similarly, what I've written above as "$\forall p^t \neg p$" would more properly be written as $\forall \neg$ "it is not the case that is t-universally instantiated". More complicated quantificational claims require the λ-abstracts, but nothing here will turn on the details of how they work.

conventions about marking every type.) More generally, if a bit more abstractly, the neo-Fregean pluralist will not only hold that $\exists x^e x = x$ ("there's something"), they'll also hold that $\exists_{e \to t} FF = F$ and $\exists_t pp = p$. Indeed, the pluralist might hold that, for every σ, $\exists_\sigma xx = x$. As I understand the position, a battery of existence claims like this will be the pluralist's official version of what I've roughly paraphrased as "there are entities of different ontological categories".[4]

As you can see, the distinctive expressive resources of the higher-order language play an important role in stating pluralism exactly. As I mentioned a moment ago, this point will turn out to be key to the central challenges of the paper. But now that I've made this point, I will often ignore it, and proceed with my rough manner of speaking. It's just much easier for me to write in English (or whatever this odd dialect is that philosophers speak) than in a higher-order language, and I hope it'll be easier for you to read too. In this vein, I'll keep using "*et*-property", "*t*-proposition", "entity" and "ontological category" even though they aren't quite right.

There's a vast amount more to say here, but it's well past time to come back to the attitudes. The standard list of identified logical constants in higher-order languages is short to state; we've already seen them all. There's $\neg$ $(t \to t)$ and $\wedge$ $(t \to (t \to t))$, and, for every σ, both $=_\sigma$ and $\forall_\sigma$. But higher-order metaphysicians typically go on to identify further constants associated with the subject-matter that interests them. For instance, arguably "it's metaphysically necessary" is not definable from the usual list of constants (in defense of this view, see Bacon, 2018, and against it, Fritz, 2023; for discussion, see Dorr et al., 2021, Ch. 8). But higher-order metaphysicians tend to hold that their favored language contains a constant corresponding to this expression (which, like $\neg$, will have type $t \to t$). A similar approach is natural for the propositional attitudes: a higher-order metaphysician interested in the attitudes will hold that their favored language contains constants of the appropriate type for each propositional attitude. For belief, the most obvious type would be $t \to (e \to t)$: the relevant expression will combine with a sentence (e.g. "some rabbits can speak", of type t) to produce a predicate ("believes some rabbits can speak", type $e \to t$), which in turn combines with a name (e.g. "Alice", type e) to produce a sentence ("Alice believes some rabbits can speak" type t). Since expressions for relations like "loves" in our higher-order language also take their arguments one at a time, the view that constants corresponding to expressions for propositional attitudes have type $t \to (e \to t)$ is naturally thought of as a form

[4] The higher-order languages I'm interested in here do not have a type-neutral way of expressing quantification. This might seem odd at first, but there's a strong argument that such a notion does not make sense. If there were a quantifier of this kind, presumably we could use it to say things like "every entity is happy", where "every entity" is intended to generalize "over" all types. This generalization would have as an instance the not merely false but the obviously nonsensical "loves is happy". Since the instance is nonsense, plausibly the generalization is too.

of relationism. I'll call this view "naive higher-order relationism", or just "naive relationism" for short.

Naive relationism offers a simple account of the objects of propositional attitudes, which solves the problem for relationism with which I began this section. Higher-order metaphysicians hold that t-propositions are the denotations of sentences. In the relevant loose sense described above, they can be true or false: there is a salient property that the denotation of "some rabbits can speak" has, since rabbits can't speak, that it would not have had, if they could. So, if higher-order metaphysicians can endorse naive relationism, they would have a ready-made account of the objects of the attitudes, which explains perhaps the most striking feature of these "objects". Everyone should agree that this would be a lovely feature of the view.

3. Propositional Attitudes and Identity

Unfortunately naive relationism conflicts with core principles of higher-order logic, at least given natural judgments about propositional attitudes. The goal of this section will be to present this conflict, which is related to Frege's well-known puzzle about the substitution of coreferring names. On its own, the conflict already threatens to undermine naive relationism, and hence an important motivation for neo-Fregean pluralism itself. But it will also be the starting point for the deeper challenges that I'll come to in the section after this one.

A nice feature of having identity-symbols not just for e but for higher types as well is that they allow us to formulate claims about the connections between identity at different types in a simple and straightforward way. A natural principle describing such connections says that, if a is identical to b (for any type of "a" and "b"), then the result of applying F to a is also identical to the result of applying it to b (from now on I'll use "a" as a schematic metavariable, not as the translation of "Alice", since she'll be mostly out of our view). In symbols this would be:

Atomic Congruence $a =_\sigma b \rightarrow Fa =_\tau Fb$.[5]

The argument of this section will turn on this principle. The principle is schematic not only on the expressions a, b, and F; it is also "schematic on types": it has instances where a and b have any type σ, and F has type $\sigma \rightarrow \tau$ (for any τ). The substituends for a, b, and F could all be complex terms (terms formed from other

[5] Earlier, I only introduced $\wedge$ and $\neg$, but in what follows I'll freely use other constants from first-order logic without comment. These can be seen either as the usual metalinguistic abbreviations, or as new constants with the obvious intended interpretation.

terms). But the identification in the consequent has to join applications (or, for emphasis "atomic applications") of the substituend of F to the substituends for a and b, respectively (hence the name "Atomic Congruence").

Atomic Congruence is an attractive principle in its own right. It is also part of classical higher-order logic. A standard way of axiomatizing identity in higher-order logic uses two principles:

Reflexivity $a =_\sigma a$; and

Substitution $a =_\sigma b \rightarrow (\varphi[a//x] \rightarrow \varphi[b//x])$.

In the second principle, "$\varphi[a//x]$" indicates that every free occurrence of the substituend for "x" in the substituend for "φ" is replaced with the substituend for "a" and similarly for "$\varphi[b//x]$", and it is required that in neither case doing so causes new variables in the substituends to be bound. Atomic Congruence follows straightforwardly from Substitution and Reflexivity; it is part of this classical theory of identity. (To see this, let φ be $Fa = Fx$ in Substitution; using classical propositional logic and the instance of reflexivity $Fa = Fa$, one can then derive $a = b \rightarrow Fa = Fb$.)

In the background in what follows, I will assume what I take to be an incontestable principle about $=_t$, the claim that if t-propositions are identical, they are materially equivalent:

Material Equivalence $p =_t q \rightarrow (p \leftrightarrow q)$.

We can now state the conflict that naive relationism creates between an intuitive judgment about a class of cases on the one hand, and Atomic Congruence on the other. I'll start with one of the cases. The name "Hesperus" refers to the planet Venus, as does the name "Phosphorus". Plato may have believed that the heavenly body he saw rise in the evening and which he called "Hesperus" was not the same heavenly body as the one he saw rise in the morning and called "Phosphorus" (*Laws* 821c). Supposing this is true, it is plausible that:

> Hesperus is Phosphorus and Plato believed that Hesperus is visible in the evening, but it's not true that Plato believed that Phosphorus is visible in the evening.

To make the conflict precise, we need to translate this claim into our higher-order language. We can give a simple such translation, using "h" (type e) for "Hesperus", "p" (type e) for "Phosphorus", V (type $e \rightarrow t$) for "is visible in the evening", and B (type $t \rightarrow t$) for "Plato believed that". Using this translation, the claim becomes:

[A] $h =_e p \wedge B(Vh) \wedge \neg B(Vp)$.

For this first presentation of the argument, I'll assume that naive relationism allows us to translate "Plato believed that" as a unary sentential operator, i.e. an expression of type $t \to t$, even though this isn't true. (I'll relax this assumption in a moment; for now it just allows us to use simpler notation.) The scheme for translation that gives us the formal claim [A] also involves a number of other assumptions which go well beyond naive relationism itself. To mention a salient one, it involves assuming that the occurrences of the names "Hesperus" and "Phosphorus" are translated in the same way into the formal language regardless of whether they occur inside or outside the scope of "believe". I won't attempt to write down all of the assumptions built in to this translation scheme. I'll think of them as bundled into [A], and discuss them later when I discuss the prospects of rejecting this premise (Section 6).

The key observation of this section is that [A] is inconsistent with Atomic Congruence (and Material Equivalence), given just classical propositional logic. The argument is as follows:

[B] $h =_e p \to Vh =_t Vp$ ("If Hesperus is Phosphorus then for Hesperus to be visible in the evening just is for Phosphorus to be visible in the evening.") (Atomic Congruence)

[C] $Vh =_t Vp \to (B(Vh) \leftrightarrow B(Vp))$ ("If for Hesperus to be visible in the evening is for Phosphorus to be visible in the evening then Plato believed that Hesperus is visible in the evening if and only if Plato believed that Phosphorus is visible in the evening.") (Atomic Congruence, Material Equivalence)

[D] $h =_e p \to (B(Vh) \leftrightarrow B(Vp))$ ("So, if Hesperus is Phosphorus then Plato believed that Hesperus is visible in the evening if and only if Plato believed that Phosphorus is visible in the evening.") ([B], [C])

And, of course, [D] contradicts [A].

Getting this conflict on the table was the main goal of this section. In the rest, I'll make three observations of detail: a first and second, about how the argument can be strengthened and generalized, and a third (most importantly) about how this argument differs from more standard presentations of a related paradoxical argument. But readers who are impatient to get to the main ideas may wish to skip to the next section at this point.

First, I assumed above that "Plato believes that" can be perspicuously translated into higher-order logic as a unary sentential operator. But naive relationism doesn't say this. It says that "believe" can be translated by a constant of type $t \to (e \to t)$ (which I'll write as "B^*"). And this difference matters, since if $t \to (e \to t)$ is the type of the translation of "believe", the argument can't be run using Atomic Congruence alone. We could use that principle to derive $B^*(Vp) = B^*(Vh)$ ("To believe that Phosphorus is visible in the evening is to believe that Hesperus is visible in the evening"), but then we'd be stuck. We could not infer from there to $(B^*(Vp))(pl) = (B^*(Vh))(pl)$ (where "*pl*" is "Plato"), the claim we'd naturally use to get a version of [C].

All of this is true, but in my view it's moot. For there are at least three compelling ways of either justifying the assumption that "Plato believes" can be translated as a unary sentential operator, or running the argument without it:

(i) First, one might see the apparent difference between these two translations of "believe" as an unimportant artifact of the particular, artificial language I chose to work with here. As we saw earlier, in this "functional" language, we must choose an order for the arguments of expressions for relations: whether the lover comes first or the beloved. It's true that naive relationists translate "believe" by B^*, which *first* takes its sentential argument and then its e-type argument. But intuitively there should not be a significant philosophical difference between this view and one which instead translates "believe" by an expression which takes its e-type argument first, followed by its t-type argument (so it'd have type $e \to (t \to t)$). If we chose that intuitively equivalent translation, we *could* run the argument using only Atomic Congruence. Seen in this light, the fact that naive relationism blocks the argument seems just an accident of notation, and not a significant philosophical point.[6]

(ii) Second, we could avoid this issue altogether by motivating the formal premise [B] using expressions in English which could be translated as unary sentential operators, for instance "it is *a priori* that", "it is obvious that", "it is surprising that", or just "obviously", "surprisingly", or "probably". If the analogue of [A] is also plausible for such expressions, then we would still have an important argument against Atomic Congruence.

(iii) Third, we could circumvent the issue by appealing, in addition to Atomic Congruence and Material Equivalence, to Application Equivalence, the principle that $F = G \to (Fa \leftrightarrow Ga)$. In this version of the argument, Atomic Congruence is used to move from [B] to: $B^*(Vp) = B^*(Vh)$, and Application Equivalence (with "Plato" substituted for a) is then used to derive [D].

To reject the argument for a conflict between [A] and Atomic Congruence on the grounds that my formalization of belief ascriptions is too simplistic, one must hold that all three ways of repairing it fail. But each seems compelling, so I'll continue to use the simpler formalization below.

[6] A natural analogue of Atomic Congruence in a *relational* type theory, where expressions can take multiple arguments all at once, *would* allow us to derive $(B^*(Vp))(pl) = (B^*(Vh))(pl)$ directly. The relational formulation would be: $a = b \to R(x_1, \ldots, x_n, a, y_1, \ldots, y_m) = R(x_1, \ldots, x_n, b, y_1, \ldots, y_m)$. This principle allows us to derive $B^*(Vp, pl) = B^*(Vh, pl)$, if we let B^* have type $\langle\langle\rangle, e\rangle$ and assume that $Vp =_t Vh$. But it does not obliterate the distinction between atomic applications and non-atomic applications; for instance, this sentence could still not be directly obtained as an atomic application to h or p.

Second, the argument applies to a much broader class of views than just naive relationism. Let higher-order relationism be the view that attitude ascriptions are to be translated into the formal language by sentences of the form $(A(\hat{t}\varphi))x$, where: x, the name of the ascribee, has type e; φ, a sentence, has type t; $\hat{t}$ (pronounced " 't'-hat"), standing for the complementizer "that", has type $t \to \pi$ for some type π; and A, corresponding to the attitude verb, has type $\pi \to (e \to t)$. We can run the argument for any form of higher-order relationism, using the following instance of Atomic Congruence:

$\hat{t}$-**Congruence** $\quad Vp =_t Vh \to \hat{t}Vh =_\pi \hat{t}Vpq.$

The argument is then as above, using Atomic Congruence and Material Equivalence to derive [C].

This generalization will be especially significant for those who see a close relationship between propositional attitude *ascriptions* in natural language and the attitudes themselves. Philosophers of this stripe might reject naive relationism because it does not include a formalization of the expression "that". But they might (for instance) instead endorse *e-relationism*, according to which expressions for propositional attitudes are regimented as having type $e \to (e \to t)$, and their apparently sentential arguments are operated on by an operator corresponding to "that" or even "the proposition that" which has type $t \to e$. This position is not a version of naive relationism, but (whatever its other merits) it is still a version of higher-order relationism, and it is thus still subject to the argument above.[7]

Third, and finally: the argument I've given differs crucially from a more traditional presentation of related ideas. This more traditional presentation turns on using the expression "is such that" to make a sentence into a complex predicate, for instance, creating "is such that Plato believed it was visible in the evening" from "Plato believed it was visible in the evening". We can translate this use of "is such that" into our higher-order languages using Church's λ-notation, rendering "is such that Plato believed it was visible in the evening" for instance as "$\lambda x^e.B(Vx)$". Here the variable after the λ (in the "head") indicates that occurrences of this variable in the "body" of the λ-term (i.e. "$B(Vx)$") are "bound". The whole predicate will have type $e \to t$: e, from the type of the variable in the "head", and t from the type of the body. In general, for any variable x of type σ, and any expression φ of type τ, the expression $\lambda x^\sigma.\varphi$ has type $\sigma \to \tau$.[8]

[7] A view of this kind fits with a popular view in linguistics, on which the English "that" has (roughly) type $t \to (e \to t)$, so that "that"-clauses have type $e \to t$, and, when "that"-clauses are used as complements of attitude verbs, they modify covert arguments which have type e (Kratzer, 2006, Moulton, 2009, Moulton, 2015, cf. Moltmann, 2003, Matthews, 2020 van Elswyk, 2019, van Elswyk, 2020, van Elswyk, 2022, Güngör, 2022; and for a slightly different view, which comes to much the same, Pietroski, 2000, Pietroski, 2005, Forbes, 2006, and especially Forbes, 2018).

[8] As promised, these expressions allow us to treat quantifiers as constants; see above n. 3. In particular, officially $\forall_\sigma x^\sigma \varphi^t$ can be understood as an abbreviation of $\forall_\sigma(\lambda x^\sigma.\varphi^t)$.

Using this notation, the more traditional argument runs as follows:

1. $h = p \to ((\lambda x.B(Vx))h \leftrightarrow (\lambda x.B(Vx))p$ ("If Hesperus is Phosphorus, then Hesperus is such that Plato believed it was visible in the evening if and only if Phosphorus is such that Plato believed it was visible in the evening.") (Atomic Congruence, Material Equivalence)
2. $((\lambda x.B(Vx))h \leftrightarrow (\lambda x.B(Vx))p) \to (B(Vh) \leftrightarrow B(Vp))$ ("If Hesperus is such that Plato believed it was visible in the evening if and only if Phosphorus is such that Plato believed it was visible in the evening, then Plato believed Hesperus was visible in the evening if and only if Plato believed Phosphorus was visible in the evening.") (Assumption)
3. $h = p \to (B(Vh) \leftrightarrow B(Vp))$ ("So, if Hesperus is Phosphorus, Plato believed Hesperus was visible in the evening if and only if Plato believed Phosphorus was visible in the evening.") (1, 2)

Arguably the most prominent responses to this argument (or something like it) have focused on premise (2) (or something like it). For instance, following Quine (1956) and especially Kaplan (1968) (cf. Kaplan, 1986), many philosophers have held that variables behave differently from names in the scope of attitude verbs, in such a way that (2) is rejected (for recent developments, see Yalcin, 2015, Lederman, 2022). But, crucially, these strategies do not offer any hope for responding to the earlier argument from [B] to [D], because that argument does not use any premise analogous to (2). It does not appeal to claims about material equivalence or identity when a term is moved from outside the scope of an attitude verb, to inside its scope. In fact, the argument does not use λ-binders or the English "such that" at all. (For further discussion, see Caie et al., 2020, pp. 6–8.)

4. Attitudes and Logic

Atomic Congruence, naive relationism, and [A] cannot all be true. In the next three sections, I'll consider the prospects for rejecting each of these in turn. In the end, I'll suggest that none of these options is attractive.

In this section, I'll focus on Atomic Congruence. As I said above, Atomic Congruence is an attractive principle in its own right, which is, moreover, a part of classical higher-order logic. Everyone should want to accept it. But one might think that propositional attitudes create pressure for everyone to reject something like this principle, so that it couldn't be the basis for a special challenge for higher-order metaphysics. In this section, I will argue that this isn't right. I'll argue that neo-Fregean pluralists have distinctive reasons to endorse Atomic Congruence, which go beyond those that monists have.

Those reasons stem from the "foundational challenge" it's the aim of this paper to present. There are in fact two of these challenges: one, metasemantic, and one epistemological. Each of them can be thought of as a dilemma, only one horn of which is directly relevant to Atomic Congruence. But I'll start with an outline of the dilemmas as a whole, before turning to how they bear on Atomic Congruence.

I'll first outline the metasemantic challenge, since, although I think it's less challenging, it's easier to get into.

As we saw in Section 2, neo-Fregean pluralists can only articulate their central commitments using distinctive higher-order languages. As a result, if key expressions of these languages are meaningless, or if they are extremely vague, the statement of neo-Fregean pluralism itself will be meaningless, or extremely vague.

One way of justifying the claim that higher-order languages are meaningful and not extremely vague would be to adduce a translation between them and some antecedently understood language (whether English or some scientific language like the language of fundamental physics, or even cognitive science). But higher-orderists can't use this strategy, because key expressions in higher-order languages cannot be translated directly into such antecedently understood languages. There is no expression in these languages that is synonymous with $=_{(t\to t)\to t}$, which joins two predicates of unary sentential operators to make a sentence, or for that matter with $\forall_{(t\to t)\to t}$, which allows quantification into the syntactic position of predicates of unary sentential operators.

Since they can't appeal to such a translation, how might the higher-orderist justify the claim that their language is meaningful and not extremely vague?[9] The best case I know of rests in part on two key ideas (for these points see Fritz et al., 2021, §5.2, Fritz, 2023). First, when we speak of the meaningfulness of higher-order languages, we mean such languages *as endowed with a strong logic*. The laws and inference rules of this strong logic winnow down the possible meanings of logical expressions in the language, helping to guarantee that these logical expressions are meaningful and not extremely vague. (From now on, I'll just say "meaningful" to mean "meaningful and not extremely vague".) Second, the meanings of expressions in higher-order languages are in part fixed by their relationship to an antecedently understood (natural or scientific) language. Even if higher-order languages cannot be translated fully into an antecedently understood language, there are *some* synonyms or at least close synonyms between higher-order languages and antecedently understood ones. For instance, expressions like "and", and sentences like "it is metaphysically necessary that either Socrates is Socrates or Socrates is not Socrates" are supposed to have more or less exact

[9] As many have observed, but bears repeating, a language may be meaningful even if it cannot be translated into another antecedently understood language. Following Frege, Prior (1971), and Williamson (2003, pp. 416–417, 459), higher-order metaphysicians have rightly observed that, even if there is no such translation, higher-order languages can still be learned by "the direct method", the same method infants use to learn their first natural language. But it's equally important to note that on its own the existence of the direct method doesn't show that seemingly learnable languages are meaningful and not extremely vague. The mere fact that a language *seems* meaningful does not show that it is. Philosophers have a terrible track record of mistaking seeming-intelligibility for the real deal. Given this, higher-order metaphysicians should have independent reasons (apart from the seeming) for thinking that their language really is meaningful.

synonyms in our higher-order language. The meanings of some expressions in higher-order languages can then be fixed by these anchors in other languages. The more anchors there are, the more there are fixed meanings for expressions in higher-order languages, and the stronger the case that the languages as a whole are meaningful.[10]

The metasemantic version of my foundational challenge is this: propositional attitudes seem to require weakening one or the other of these two pillars, and so to require abandoning this best case for the meaningfulness of higher-order languages. If naive (or, more generally, higher-order) relationism is accepted, then given [A], the pluralist is forced to give up Atomic Congruence, thereby weakening the first pillar, of a strong logic. And if naive (or, more generally, higher-order) relationism is rejected, the pluralist loses out on a rich and important class of anchors for their higher-order language, thereby weakening the second pillar. (Or, at any rate, that's what I'll argue in the next section.)

There is an obvious response to this metasemantic challenge, which will help to motivate the second, epistemological challenge. Many philosophers hold that expressions can have determinate meanings even without many relevant "anchors". A prominent way of articulating this point appeals to the notions of "naturalness" and "magnetism". On a standard view, some meanings are so "natural" that they have a special kind of "magnetism" which "attracts" some expressions to them. Those who endorse a view of this kind can readily dismiss the second horn of my metasemantic challenge. For them, a rich class of anchors is comparatively unimportant. If reality is the way the neo-Fregean says it is, then the denotations of (say) $\forall_t$ or $=_t$ will be extremely natural, and the magnetism of these meanings will guarantee that the expressions are meaningful (and determinately so).

Whatever its merits, this response does not escape the second, *epistemological* version of the foundational challenge. The two pillars of a strong logic and rich anchors in an antecedently understood language are not just important evidence for the meaningfulness of higher-order languages; they are also evidence for the metaphysical vision the pluralist uses those languages to articulate. The consistency and strength of neo-Fregean pluralism, embodied in classical higher-order logic, is part of our (abductive) evidence that neo-Fregean pluralism is true. Moreover, natural language is the basis for our first exposure to the array of "ontological categories" that neo-Fregean pluralists claim there are. If higher-order languages are cut loose from anchors in natural language (or from some other

[10] A third idea emphasized in Fritz et al. (2021) is that there are analogies between certain constants at different types. For instance, we can understand the universal quantifier at higher types by analogy to the universal quantifier at lower types. A fourth is that the theoretical utility of the language may help provide evidence for its meaningfulness. I believe that these ideas are as important as the others, but for simplicity in the main text I'll just talk about the other two. Fritz, who is responsible for the ideas in Fritz et al. (2021, §5.2), develops a metasemantic account based on these "pillars" in more detail in Fritz (2023).

antecedently understood languages), we lose a key source of evidence for these "ontological categories", and (as I'll discuss more in a moment) we also lose out on key sources of evidence for the character of the entities which belong to them.

Each of these foundational challenges give higher-order metaphysicians special reason to accept Atomic Congruence. First, higher-order metaphysicians introduce a new language to articulate their view, and must justify the claim that this language is meaningful. And second, higher-order metaphysicians must provide evidence for their distinctive ontology. Endorsing a strong logic, most obviously, some extension of classical higher-order logic (which contains Atomic Congruence), will help to meet both of these challenges. These are distinctive challenges for the higher-order metaphysician, which need not arise for objectual monists in general, since monists do not have to understand their key terms as untranslatable into antecedently understood languages, and do not endorse the same distinctive ontology the higher-orderist does.

These points apply generically to the principles of higher-order logic and thus support Atomic Congruence. But there are in addition further, specific reasons to endorse Atomic Congruence, which are in a way more interesting, and more powerful. Let's start again with metasemantics. Atomic Congruence helps to constrain the meanings of expressions in higher-order languages in at least three three different ways:

(i) It constrains the meaning of $=_\sigma$, for all σ, by constraining how identity statements at one type relate to identity statements at other types.

(ii) It is attractive to think that expressions of type e and t play a special metasemantic role in fixing the meanings of expressions of all types. Given the structure of the type-hierarchy, one might think that if expressions of these types have determinate meaning, then expressions of all types do. But this attractive idea plausibly depends in part on endorsing Atomic Congruence. For instance, $a =_e b \rightarrow Fa =_t Fb$, provides important constraints, not just on $=_e$, and $=_t$, but also on the meanings of predicate expressions which can be substituted for F; such expressions cannot "cut more finely" than $=_t$. Given Atomic Congruence, a similar point can be extended from the "base types" e and t to all types. Without Atomic Congruence, we cannot tell this simple story.

(iii) Together with Material Equivalence, Atomic Congruence implies:

Atomic Substitution $a =_\sigma b \rightarrow (Fa \leftrightarrow Fb)$.

This principle says that the truth and falsity of atomic applications constrain the meaning of identity at all types. If Atomic Congruence and Material Equivalence can be accepted together with higher-order relationism, then the truth and falsity of propositional attitude ascriptions (whether in English or a formal, scientific language) in particular would provide important constraints on $=_t$.

Just as in the metasemantic case, there is also an epistemological argument which specifically targets Atomic Congruence. If a theory implies both that there is a basic question about the universe, and that that question cannot be settled, then that is typically some evidence against the theory, by contrast to competitors which do not have these implications. For instance, one version of the view that space is absolute implies that there is a basic question about where we are located in this absolute space, while also implying that we could not in any normal way determine the answer to this question. These facts provide evidence against this simplistic view. My suggestion is that higher-order metaphysicians who reject Atomic Congruence would place themselves in something of a similar position. According to higher-order metaphysicians, questions about the "fineness of grain" of properties and propositions, for instance whether in general $p =_t (p \wedge (q \vee \neg q))$, are central, basic questions about the universe. In this vein, Jeremy Goodman writes: "How fine grained is reality? This is perhaps the deepest question in all of metaphysics, and higher-order languages provide the tools to precisely formulate and productively debate competing answers to it", "allow[ing] us to ask after general principles governing a dyadic sentential operator . . . analogous to the first-order identity predicate" (Goodman (2017, 51–52)). If Atomic Congruence is rejected we lose an important tool—arguably, our only tool—for answering these deep questions. As in (i) above, claims about identity and distinctness at type τ no longer provide evidence about identity and distinctness at type σ. As in (ii), we cannot hope to understand higher-order identity as fixed by identity at types *e* and *t*. And, most importantly, as in (iii), if Atomic Substitution is rejected in addition to Atomic Congruence, we lose the possibility of appealing to judgments of truth and falsity of atomic predications with sentential arguments to constrain identity among *t*-propositions.

The resulting situation is certainly not exactly the same as the one I described above for the simplistic theory of absolute space. Higher-order metaphysics without Atomic Congruence does not imply that there is *no* possible evidence we could get to help us answer these basic questions. But still, giving up on Atomic Congruence brings the neo-Fregean pluralist's view of propositional identity closer to this view of our locations, and to that extent it provides evidence against this form of pluralism.

5. Attitudes and Anchors

Neo-Fregean pluralists thus have special reason to uphold Atomic Congruence. In this section I will consider the prospects of responding to the argument of

Section 3 in a different way: by rejecting naive relationism (and, more generally, higher-order relationism). I'll argue that this option too comes at a high price.

As I have said, the possibility of giving an account of the "objects" of propositional attitudes has historically been an important motivation for neo-Fregean pluralism. But a prominent strain of neo-Fregean pluralists today do not see this motivation as central to the program, and indeed reject the claim that *t*-propositions stand in a straightforward relationship to the objects of attitudes. Cian Dorr (2016) explicitly cautions against understanding the referents of sentences in his favored higher-order language as closely connected to propositions construed as the objects of propositional attitudes, precisely on the grounds that doing so would lead to what he sees as absurd results about the distinctness of *t*-propositions, for instance that $Vh \neq_t Vp$. Dorr suggests instead that *t*-propositions should be understood as "entities which stand to the number 0 as properties stand to the number 1 and binary relations stand to the number 2" (Dorr, 2016, p. 54; see now Dorr et al., 2021, p. 6, n. 5; for more discussion, see also Menzel, this volume, pp. 432–433). And Jeremy Goodman presents a related view, using the terms "representation" and "reality" to distinguish, roughly, between propositions as the objects of attitudes and *t*-propositions as he understands them, which might here be closer to "states of affairs" (Goodman, 2017, 51–52, cf. Fritz, 2017, p. 56).

Without characterizing it exactly, we might call this orientation toward the subject matter of quantification into sentence position *mondialism*. Mondialism begins with a "worldly" conception of *t*-propositions, and does not assume at the start that they are closely related to the objects of propositional attitudes. This orientation contrasts with the sort of view we've been exploring until now, and which we might call *idealism*: an orientation which starts from the idea that *t*-propositions just are the objects of propositional attitudes.

Mondialism does not require rejecting higher-order relationism. One might start with a mondialist orientation, but conclude that, in the end, even naive relationism is correct: the attitudes do relate us to these worldly entities. But unlike idealism, mondialism does not have naive relationism baked into it from the start. A mondialist is free, for instance, to endorse a form of "sententialism", or "quotationalism", according to which attitudes are relations between people and sentences, not between people and *t*-propositions, or even between people and the result of applying some operator to a *t*-proposition (e.g. Quine, 1956, Quine, 1960, Davidson, 1968, Field, 2001, Field, 2017). Such a position makes it natural to reject both [A] and [C] in the argument of Section 3, on the grounds that "believe" should be translated as a predicate of sentences, not as a monadic sentential operator. And sententialism is not the only view which could drive this response. Mondialists could similarly reject higher-order relationism if they hold that attitudes are relations to something like Fregean thoughts, that *t*-propositions are analogous to what those thoughts present, and that there is no "backward road" from what a thought presents to the thought itself.

But whatever the virtues of these views, I will now argue that, because they reject higher-order relationism, they face the second horn of the two foundational challenges introduced in the previous section.

First, there is the metasemantic one. Any view which rejects higher-order relationism to some extent requires weakening the second pillar of the best case for the meaningfulness of higher-order languages, since it requires abandoning an important class of anchors for expressions in the higher-order language. Other things being equal an expression faces less of a threat of being meaningless or extremely vague if its meaning is constrained by a large and varied class of uses either of it or of its synonyms. If sentences of English or an antecedently understood scientific language can be translated directly as sentences in the higher-order language (i.e. they are synonyms), then those sentences in the higher-order language will be constrained by a large and varied class of uses. And the fact that these sentences are well-anchored will, in turn, help to constrain the meaning of the quantifier $\forall_t$ since it constrains the meanings of sentences which count as instances of generalizations formed using it.

This story does not specifically mention *embedded* sentences, and one might wonder what would be lost if we left them out altogether. Wouldn't it be enough if (say) every unembedded sentence of English were translated into a sentence of the higher-order language, even if embedded sentences (for instance in the complements of attitude verbs) were not? But giving up on translating embedded sentences as sentences would come at a significant cost. The reason is directly related to the specific metasemantic role of Atomic Congruence described in the previous section, most importantly point (iii). To recall, I said there that Atomic Congruence (together with Material Equivalence) implies Atomic Substitution, i.e. the principle $a =_\sigma b \rightarrow (Fa \leftrightarrow Fb)$, and that this principle in turn ensures that the truth and falsity of atomic applications constrains the meaning of identity at all types, and in particular at type t. The constraints thus imposed on $=_t$ in particular are stronger to the extent that there is a richer class of sentences of the form Fa (where a has type t) whose truth and falsity are known. If attitude ascriptions are included in this class of atomic applications, then the truth and falsity of these ascriptions provide a rich class of constraints on $=_t$. If they are not, we lose out on such constraints. And since, as we've seen (point (ii) above), fixing the meanings of expressions of type e and t may help to fix those of all higher types as well, these constraints on type t may be of special importance.

This is the crux of the metasemantic dilemma: as we saw in the previous section, rejecting Atomic Congruence means losing important constraints on $=_t$. But (as we've seen here) accepting Atomic Congruence and rejecting higher-order relationism also means losing important constraints on $=_t$, for essentially the same reason.

As I've discussed, however, there is a fairly well-understood response to the metasemantic challenge: to hold that the meanings of higher-order identity or the higher-order quantifiers are distinctive ("natural", "magnetic") in such a way that these expressions are determinately meaningful without having a rich class of anchors. But this line of thought does not offer a response to the second, *epistemological* challenge, namely, that neo-Fregean pluralists who reject higher-order relationism give up on what would otherwise be important evidence about the existence and character of t-propositions. If natural language or some other antecedently understood theoretical language systematically draws distinctions similar to those drawn in higher-order languages, this would be at least some evidence for the distinctions higher-order metaphysicians claim there are. But if the objects of propositional attitudes are not t-propositions (or something closely related to them), then these distinctions are relevantly different, eliminating one line of support for the higher-orderist's ontology. (A version of this concern may be particularly pressing for those who endorse the response to the metasemantic challenge just mentioned, since eliminating embedded sentences would also eliminate evidence that the class of t-propositions is "natural".)

This version of the epistemological challenge concerns our evidence for the existence of t-propositions. But a more challenging version of it focuses on our evidence about their character. At the end of the previous section, I said that, if Atomic Congruence is rejected, it is hard to see even in principle how we could answer questions about the fineness of grain of properties and propositions, which higher-orderists deem to be deep ones about the nature of reality. I suggested that this fact would provide some evidence against higher-order metaphysics. But note that a similar point would still apply even if Atomic Congruence is accepted, but higher-order relationism is rejected. In the extreme case, suppose that a higher-orderist accepts Atomic Congruence but denies that any notions we are familiar with give us a handle on expressions of type $t \to \sigma$ for all types σ. Here, accepting the letter of Atomic Congruence clearly does little to change the basic epistemological situation I discussed at the end of the last section: since there are no instances of Atomic Congruence that we can assess directly, this principle does not help to provide evidence about the identity and distinctness of t-propositions in the same way it otherwise might. The case of a mondialist who rejects higher-order relationism is of course less clear-cut than this extreme position. But arguably it is still similar in key respects. For this mondialist, Boolean connectives like "and" or "not" and alethic modals like "it's necessarily the case that" do provide us with instances of Atomic Congruence, while propositional attitudes do not. But on their own (if propositional attitudes are not considered), Booleans and alethic modals give us quite limited constraints on theories of higher-order identity. Simplicity and strength may provide further evidence in favor of one or another theory of propositional fineness of grain, but these constraints together will still leave us with at best a bewildering array of reasonable theories of fineness of grain.[11]

[11] Grounding might help, but so far most of what we have are inconsistency results Fritz (2022), Fritz (2023), Wilhelm (2021), though cf. Litland (2022) and Goodman (2023).

This problem for mondialists is not as clear-cut as the one that the toy theory of absolute location faced in my description at the end of the last section. My bet is that the behavior of Boolean connectives and alethic modals combined with simplicity and strength on their own will not give us answers to questions about fineness of grain. Slightly more speculatively, my guess is that there may not be any good methods to answer those questions if propositional attitudes are discounted. But this bet and guess are obviously less certain than the claim that the toy theory of absolute location entailed that no normal methods could help us to determine where we are. Still, I suggest that, since rejecting higher-order relationism brings neo-Fregean pluralism closer to theories of absolute location, if pluralists must reject higher-order relationism, this would be some evidence against the view.

In writing this paper, I've vacillated a great deal on how serious this last point is. Some days I think it's very weak: sure, pluralism would be in better shape if it were compatible with higher-order relationism, but that's not to say it's in bad shape if it's not. Other days I think it's strong: it captures the sense in which, if propositional attitudes are discounted, we don't have a good grip at all on what sentences in the higher-order language denote, or what $\forall_t$ quantifies "over". I'll assume in what follows that the strong reading is the right one, but that's not because I'm convinced of it. I hope this section will serve as an invitation for others to arrive at a more definite verdict here, whether by showing that there is no problem, after all, or by making the problem more exact than I've been able to.

6. Substitutionalism

This leaves us with one final response to the argument of Section 3 to consider: that of denying [A]. I'll focus my discussion on *Substitutionalism*, the family of views which reject [A] and uphold not just Atomic Congruence but full-blown Substitution. (I'll call examples like [A], which conflict with Substitution, cases of *opacity*; those who accept that there are such cases are *opacitists*. I discuss some forms of opacitism in more detail in the Appendix.)

The core arguments of this paper are the ones I've developed in the previous two sections, against rejecting Atomic Congruence on the one hand, and against rejecting higher-order relationism on the other hand. So I would consider the paper a success if the moral readers take away is that higher-order metaphysicians have special reasons to endorse both Atomic Congruence (and, in that case, plausibly, Substitution) and higher-order (or even naive) relationism.[12] But I myself am not convinced that this is the correct conclusion to draw. I agree that

[12] Late in the writing of this paper, I found Trueman (2022), which develops an argument similar to the one I presented in Section 3, and endorses essentially this response to it.

Substitutionalism is attractive if we consider just the problems I've been focused on here. But Substitutionalism has a slew of other well-known problems. These problems are hard enough that they suggest a stronger result, namely, that if higher-order metaphysicians have special reason to endorse Substitutionalism, that is some reason to reject higher-order metaphysics itself.

To give a feeling for why one might be driven to this stronger conclusion, I want to just sketch some arguments against a few different versions of Substitutionalism. There's an enormous literature on positions like Substitutionalism in an objectual monist setting, and I can't even begin to discuss all of the issues here. But I hope the challenges and views I'll discuss are at least somewhat illustrative of the general shape such a discussion might take.

A first style of Substitutionalist view accepts that something like [A] is the correct translation of the informal sentence which motivates it, but simply rejects the claim made by that informal sentence. This kind of *hardline Substitutionalism* faces several important objections. A first is that the resulting picture of mental content is unintuitive, since there certainly seems to be a difference in *content* between the belief that Hesperus is bright and the belief that Phosphorus is bright. Second, it is unclear whether such a theory of mental content can be the basis for adequate explanations of behavior (see e.g. Heck, 2012, with Almotahari and Gray, 2021; cf. Chalmers, 2011, Braun, 2016, Chalmers, 2016). Third, hardliners must not just adopt an unintuitive position for examples of the sort described in [A]; they must also explain other cases of apparently identical *t*-propositions which are expressed by sentences that do not appear to be interchangeable in the context of attitudes. For instance, many think that sentences which express logical truths express the same proposition, but it seems that people may (for example) believe that if rabbits can talk then rabbits can talk, without believing that if *PA*2 then *FLT* (where *PA*2 is the conjunction of the axioms of second-order Peano arithmetic, and *FLT* is the statement of Fermat's last theorem). On some relevant theories of propositional identity, this pattern would directly give us a counterexample to Atomic Congruence, so the hardliner would have to reject the intuitive judgment here. Different theories of propositional identity might allow us to escape this particular example, but most reasonable theories I'm aware of predict that there are some intuitive apparent counterexamples to Atomic Congruence (an exception is Fritz et al. 2021). So in general hardline substitutionalists must not just take a hard line with [A]; they must (less plausibly) do so with these further examples as well. Fourth, and finally, this kind of view faces steep challenges in making sense of English attitude ascriptions. If English is taken to be a language in which Substitution fails, there is a question about how to map English systematically to the well-behaved theoretical language for which Substitution holds (see e.g. Yli-Vakkuri and Hawthorne, 2021). If on the other hand [A] is simply claimed to be false in English, a great deal needs to be said to make this astonishing claim remotely plausible. I hope it's fair to say that the existing defenses of this claim, although philosophically rich, have not won wide acceptance (see e.g. Salmon, 1986, Soames, 1987, Braun, 1998, Saul, 2010, Williamson, 2021, Williamson, 2021).

A second style of view accepts the English version of [A] but rejects the way I've translated it into the formal language.[13] According to the version of this view I'll consider in detail, the informal sentence which motivates [A] receives a true reading only if one or more expressions in it are interpreted ambiguously; the formalized [A] itself fails to reflect this non-uniform interpretation of relevant expressions. On one version of this view the verb "believe" is context-sensitive and interpreted differently in its different occurrences (Loar, 1972, Schiffer, 1979, Crimmins and Perry, 1989, Percus and Sauerland, 2003, Dorr, 2014, Goodman and Lederman, 2021); on another, the names themselves can be translated by expressions of different types depending on where they occur in the sentence (Thomason, 1980, Muskens, 1991, Muskens, 2004); and there are many more besides. These proposals all promise to help with the fourth problem for hardline Substitutionalists: that of giving a reasonable theory of English ascriptions. In the end it's not clear how successful such views are even on that front, but here I won't litigate this point (I've discussed some challenges in Goodman and Lederman, 2021, §9, Lederman, 2021). What I want to focus on instead is that, even supposing they do offer adequate theories of English, such theories seem to make little progress at all on the first three problems for hardliners. Some authors claim not to believe that propositional attitudes are anything other than the relations expressed by natural language sentences featuring propositional attitude verbs. But many reject this view; they hold that the propositional attitudes are important features of the mind which aren't defined by the idiosyncracies of this or that natural language expression.[14] On this view, while there may well be context-sensitivity or ambiguity in the way we talk about the attitudes *in English*, there presumably isn't context-sensitivity or ambiguity in a better-regimented language which allows us

[13] Yli-Vakkuri and Hawthorne (2021) offer a sophisticated discussion of possible translations which map logically deviant English reports into a better behaved higher-order language (in which Atomic Congruence holds). These proposals fit into the present taxonomy in different ways depending on whether we understand propositional attitudes to be the relations expressed in the formal analogue of English (in which case Atomic Congruence and also the analogue of (2) in the argument at the end of Section 3 will fail), or as the relations in the well-behaved higher-order language (in which case [A] will fail). The second way of taxonomizing the views they consider is most relevant for our purposes, since then their paper offers a rich study of how we might (a) uphold naive relationism, while (b) denying [A] and yet still (c) offering a systematic rendition of the poorly behaved analogue of English into the higher-order language. Hawthorne and Yli-Vakkuri show that a number of options for these translations are not particularly attractive, and conclude that probably [A] should be rejected even in English. This conclusion is broadly in line with the upshot of the arguments in the previous two sections.

[14] This question has of course been well discussed by many; for my own thoughts on a kind of middle way, see Blumberg and Lederman (2020, §8).

to talk more directly about the attitudes themselves. So the views under discussion may not get to the heart of the challenge. Even if they help us to respond to the fourth objection in the previous paragraph (helping to give a plausible theory of English), and and may help a bit with the first (if we understand our intuitions about mental content to be based on judgments of English sentences), they don't seem to help with the second and the third objections there.

As I've said, this is nothing like a full accounting of issues raised by Substitutionalism, or of possible responses to them. I would consider the paper a success if readers conclude from it that higher-order metaphysicians should be Substitutionalists. But this shouldn't be the end of the story, since Substitutionalism itself faces hard and important objections.

7. Conclusion

Higher-order metaphysicians have reason to endorse higher-order relationism about propositional attitudes, since doing so allows them to anchor their talk of *t*-propositions more firmly in antecedently understood languages, and also to find evidence for their sweeping vision of reality in the distinctions drawn by these languages. But this is harder than it might seem, since endorsing higher-order relationism forces higher-order metaphysicians either to give up on a classical logic of identity (and in particular Atomic Congruence) or to adopt a radical view about the content of propositional attitudes (Substitutionalism).

There has recently been an explosion of work on the "fineness of grain" of propositions and properties, inspired in part by the fact that so-called "structured" theories of propositions are inconsistent in higher-order logic (Dorr, 2016, Goodman, 2017, Bacon, 2020, Fritz, 2023, Fritz et al., 2021, Uzquiano, 2022, Bacon, 2023, cf. also Goodman, 2018). As I mentioned in Section 5, much of this work has been conducted from a "mondialist" standpoint, without taking a stand on how such theories of propositions might impact our understanding of propositional attitudes in particular. But some of the most prominent historical motivations for structured theories propositions derive from concerns that coarse-grained theories of propositional identity may fail to account for our judgments about propositional attitudes and rational inference (see e.g. Carnap, 1947, §13, Soames, 1987). We do not yet have a full understanding of how the new theories of the fineness of grain of propositions do or do not help with these concerns (there's some discussion in Goodman and Lederman, 2021; Caie et al., 2020; Fritz et al., 2021). Any attempt to achieve such a full understanding will require engaging with the core issues raised in this paper, and in particular whether propositional attitudes do or do not provide evidence about the character of *t*-propositions.

Although, as I've said, many questions remain open here, higher-order metaphysics does offer a distinctive perspective on one issue which has been central to objectual monists' discussions of related issues. In contemporary metaphysics and philosophy of language, questions related to "logical omniscience" are often grouped together under the label of "hyperintensionality" (see e.g. Berto and Nolan, 2021). Users of this label often hold that the distinction between the "intensional" (which can be modeled using possible worlds) and the "hyperintensional" (which cannot) marks a key watershed between intuitively well-behaved phenomena on the one side, and intuitively poorly behaved phenomena on the other. But from the perspective of higher-order metaphysics, the idea that this is an important watershed of well-behavedness is not quite right. There are strong, simple, well-behaved higher-order theories of propositional identity which reject key identities entailed by intensional theories of propositions, for example, the claim that $\varphi =_t (\varphi \wedge (\psi \vee \neg\psi))$ (see Dorr, 2016, §8, Goodman, 2018, as developed in Caie et al., 2020, Appendix C2.2, Bacon, 2020, Fritz et al., 2021). These theories draw hyperintensional distinctions but there is no special challenge in formulating them; they certainly do not have the aura of mystery that often accompanies the use of the term "hyperintensionality". Some have argued that intensionalism—roughly, the view that $\varphi =_t \psi$ is true whenever it is metaphysically necessary that $(\varphi \leftrightarrow \psi)$—or at least Booleanism—roughly, the view that $\varphi =_t \psi$ is true whenever $\varphi \leftrightarrow \psi$ is a theorem of propositional logic—is attractive on the grounds that it is in some sense the simplest credible theory of propositional identity (see Fritz, 2023, and Bacon and Dorr, this volume). But these arguments do not involve attributing incoherence or even lack of rigor to proponents of alternative theories which draw hyperintensional distinctions.

From the higher-order perspective, the line between the intensional and the hyperintensional is not one between what is coherently (or even easily) statable and what is not, this boundary falls instead between theories which endorse Substitution, and those which do not (see the Appendix for discussion).[15] Distinctions that are finer-grained than the "intensional" present no in principle challenge to higher-order metaphysicians. But phenomena (if there are any) which draw distinctions finer than identity do. For the higher-order metaphysician, we might say, opacity is the new hyperintensionality.

Will post-Quinean monism come to look to us as Milesian monism looked to the later Greeks? Or will the fervor behind neo-Fregean pluralism, with its irreducible kinds of entities, burn itself out? Nothing I've said here will settle the debate. But I hope to have made the case that propositional attitudes raise foundational questions for higher-order metaphysics, which should not be shunted to one side as matters "merely" to do with representation.

[15] In the (by contemporary lights, confusing) terminology of Kaplan, 1964, pp. 16–20, the distinction is between "extensional" and "non-extensional" systems (see the Klement, 2001, pp. 98–101).

APPENDIX: OPACITISM

This appendix discusses how the central challenges of this paper do or do not apply to various *opacitist* views, according to which there are counterexamples to Substitution ("cases of opacity"), and in particular (for our purposes here) according to which [A] and claims like it are such counterexamples. All of the opacitists I will be concerned with accept [A], so they all hold that e-Substitution has false instances. (I'll use σ-Substitution as a name for Substitution restricted to instances where the substituends for a and b have type σ.) The argument in Section 3 raises the question of whether opacitists will also allow failures of t-Substitution. Opacitists who accept [A] and [B] but reject [C] are committed to there being such failures. Any position like this is a form of *wide opacitism*; "wide" because it countenances failures of Substitution at more types then e. Those who reject [B] but accept [C] endorse instead a form of *deep opacitism*; "deep" because it allows that opacity can arise in a signature which contains only distinguished vocabulary (like $=_t$) and "intuitively extensional" expressions (like "is visible in the evening") (e.g.: $h =_e p \wedge (Vh =_t Vh) \wedge \neg(Vh =_t Vp)$).

Wide Opacitism

Wide opacitism has received some significant recent attention. But I will argue here that, from the perspective of the foundational questions I've been exploring, deep opacitism has important advantages.

I'll focus on wide opacitists who are naive relationists.[16] These wide opacitists accept [A] and [B], but reject [C]. They thus accept:

[A_t] $Vh = Vp \wedge B(Vh) \wedge \neg B(Vp)$ ("Hesperus is visible in the evening = Phosphorus is visible in the evening and Plato believed that Hesperus is visible in the evening, but did not believe that Phosphorus is visible in the evening.")

Now consider the principle:

Universal Instantiation $\forall x\varphi \to \varphi[\alpha/x]$, provided x is free for α in φ.

Given Universal Instantiation, [A_t] is incompatible with both

[16] My targets in this section will be naive relationists, but I now think that $\hat{t}$-relationism may offer a more attractive version of wide opacitism. One reason is that $\hat{t}$-relationism (but not naive relationism) is consistent with holding that "believe" (and other attitudes) are not themselves opaque (only $\hat{t}$ is). This is of interest in the context of theories which hold that the only relations which exist are transparent ones; in such a setting, wide opacitists who are naive relationists would have to deny that the propositional attitudes exist, which is a significant cost, especially for those who wish to offer broadly "functionalist" treatments of those attitudes. For discussion, see Caie et al. (2020), text surrounding n. 50.

Quantified Substitution $\forall x \forall y(x = y \rightarrow (\varphi \leftrightarrow \varphi[y/x]))$; and

Leibniz's Law $\forall x \forall y(x = y \rightarrow \forall X(Xx \leftrightarrow Xy))$.

This fact is the basis for an important argument against wide opacitism. Universal Instantiation is a key component of the logic of quantification, and both Quantified Substitution and Leibniz's Law are key components of the logic of identity. Regardless of whether wide opacitists reject Universal Instantiation (following Bacon and Russell, 2019, Bacon, 2021) or Quantified Substitution and Leibniz's Law (Caie et al., 2020), they must abandon a classical logic of identity and quantification. In particular, they must do so for type *t*, giving up principles that would otherwise help to "winnow down" the meanings of $=_t$ and $\forall_t$.

Why is this a special problem for wide opacitists? After all, all of the opacitists I'm concerned with accept [A], so a parallel problem to the one just described arises for them at type *e*: all opacitists must either reject *e*-Universal Instantiation, or *e*-Quantified Substitution, and either $e \rightarrow t$-Universal Instantiation or *e*-Leibniz's Law. (Here the modifier "σ-" indicates the type of the variable *x* in the principle.) Given this, it might seem that the considerations about quantification and identity at type *t* in the previous paragraph cannot present a problem for wide opacitists in particular, as opposed to opacitism in general.

But this response misses an important difference between types *e* and *t*, namely, that the particular metasemantic and epistemological problems I have been discussing do not arise in the same way for type *e* as they do for type *t*. (This is not to say that the metasemantics and epistemology of universal quantification or identity for type *e* are unproblematic—they aren't—it's just to say that *these* problems don't arise for them.) On a wide array of metasemantic theories, the meanings of $=_e$ and $\forall_e$ may be settled independently of whether they satisfy general logical principles like Universal Instantiation or Quantified Substitution. Many philosophers reject Universal Instantiation because of non-denoting names like "Pegasus" or because they see this as the best way of allowing that some individuals exist only contingently. While some may allege that this view about the logic of $\forall_e$ (and similarly non-classical views about $=_e$) leads to indeterminacy in the meaning of this expression, it is reasonable to reject these allegations on the grounds that it is perfectly clear what we mean by "everything" (or "the same thing"), independently of any general laws that this expression might satisfy. The discussion in this section poses a special problem for wide opacitists because this style of response to allegations about the indeterminacy of the quantifier for *t*-propositions or $=_t$ is far less appealing. It is just less clear what further data could fix the meaning of these new theoretical terms. Similarly, on the epistemological side: general logical principles governing identity and quantification are not obviously central to the question of whether there are any objects or whether I am distinct from the Eiffel Tower. By contrast, these logical principles do seem central to the evidence we have for the existence of t-propositions and whether (say) $p \neq_t (p \wedge (q \vee \neg q))$.[17]

Deep Opacitism

Deep opacitism has been much less discussed than wide opacitism. But unlike wide opacitism, versions of deep opacitism can vindicate a strong logic of $\forall_t$ and $=_t$. In fact, if deep opacitists can uphold not only t-Substitution but σ-Substitution for all $\sigma \neq e$, they can preserve a strong logic of quantification and identity at all types other than e.

In this section I'll suggest that, as least far as I can see, there may be an attractive view of this form. I'll make the case for this claim by responding to an important objection. The objection arises from cases like the following, which are sometimes (loosely) referred to as "Mates' puzzle" (Mates, 1952). Recall that"woodchuck" and "groundhog" are just two different ways of naming the same kind of animal, so that plausibly to be a woodchuck is to be a groundhog. Consider then:

Context Barbara, a monoglottal English speaker, thinks that groundhogs are blind like mole rats, while woodchucks are sighted and are often seen above ground. There is a woodchuck/groundhog who lives in Barbara's neighborhood, who is known to Barbara and other locals as "Alonzo". Barbara recognizes that Alonzo is a woodchuck, but she thinks he is not a groundhog; in fact, she believes she's never seen a groundhog:

Woodchuck To be a woodchuck is to be a groundhog, and Barbara believes that Alonzo is a woodchuck but she does not believe that Alonzo is a groundhog.

[17] For this reason I now think that there was something misleading about one key point in Caie et al. (2020, p. 17). Bacon and Russell (2019) say that, given classicism, the relation of *having all the same properties* will in a certain sense be "more identity-like" than identity because that relation would satisfy the analogue of Leibniz's Law, rendering classicism incoherent. Caie et al. (2020) point out in response that, since Hesperus and Phosphorus are identical, and classicists believe that *having all the same properties* does not relate Hesperus and Phosphorus, then clearly the latter relation is in an important sense not "identity-like". I still think this is correct. But it skates past an important issue, since at higher types we are more dependent on general logical principles to fix the meaning of identity. The insistence that we have a strong independent grasp of what counts as identity seems less compelling to me at non-e types in general than it does at type e.

We can rewrite this as $W =_{e \to t} G \wedge B(Wa) \wedge \neg B(Ga)$, where "$W$" translates "is a woodchuck", "G" "is a groundhog", "B" "Barbara believes that" and a "Alonzo". This is just one case, but the point suggested by it is a much more general one. For almost any pair of seemingly synonymous predicates, it seems we can follow this recipe to construct an example where they appear to fail to be intersubstitutable.

If deep opacitists endorse $e \to t$-Substitution, they must deny Woodchuck as formalized above. In itself, this may not be a bad position, but it is unclear whether it is open to deep opacitists in particular. After all, deep opacitists accept [A], and there may seem to be no obvious difference between [A] and Woodchuck.

But a little reflection may eliminate this appearance. Deep opacitism is motivated in part by the idea that propositional attitudes are not distinctive sources of opacity. Proponents of this view reject [B] precisely because they hope to preserve the idea that "believe" is not opaque with respect to its "objects". But if *propositional* attitudes are not opaque, then one would expect that, if there are any *propertarian* attitudes (i.e. attitudes which relate people to properties), then those attitudes too are not opaque. This natural parallelism between propositional attitudes and propertarian ones, however, pushes deep opacitists to reject Woodchuck as formalized above. For, if one accepts that Woodchuck is a counterexample to $e \to t$-Substitution, presumably "Barbara wants to be a woodchuck but does not want to be a groundhog" could also be such a counterexample. (Assuming, perhaps falsely, but for greater concreteness, that "want" expresses a propertarian attitudes.) But allowing this kind of opacity in propertarian attitudes seems to undercut a key motivation for deep opacitism, i.e. of holding that attitudes in general are not opaque with respect to their "objects".

I conclude that deep opacitists can and should endorse σ-Substitution for all $\sigma \neq e$. But this still leaves them with an important question: what should they say about Woodchuck? Two options are most salient: they might deny the first conjunct, and claim that $W \neq_{e \to t} G$; or they might deny the second or third, holding that contrary to appearances, either $\neg B(Wa)$ or $B(Ga)$. I'll discuss these responses in turn.

A deep opacitist inclined to hold that $W \neq_{e \to t} G$ might hold more generally that no two atomic predicates of type $e \to t$ express identical *et*-properties. This idea fits naturally with the deep opacitist's approach to type t: even though there is a sense in which the claim that Hesperus is visible in the evening "makes the same demands on the world" as the claim that Phosphorus is visible in the evening (for instance, necessarily the one holds if and only if the other does), the deep opacitist holds that these are distinct t-propositions. Similarly, even though there is a sense in which being a woodchuck "makes the same demands on an individual" as being a groundhog, the deep opacitist might hold that in fact these two *et*-properties are distinct.

This is an interesting position which deserves further investigation. But it faces an important challenge. Nontrivial property identities have a special place in philosophical analysis and (more importantly) scientific discovery (see Bealer, 1994). Those who accept that, for instance, to be salt is to be *NaCl* have a natural story about why features of salt can be explained by features of *NaCl*; they will say that explanation stops with identity. But if being salt is distinct from being *NaCl* it is less clear why such explanations succeed. Presumably our deep opacitists will want to say that the two are "metaphysically equivalent" in some sense, but they must then say more about this notion of metaphysical equivalence and, in particular, why it guarantees that explanations like this one succeed. A concern is that any way of answering this last question will require postulating "brute necessities".

Suppose a deep opacitist is moved by these considerations to reject the second or third conjunct of Woodchuck. What might they say to distinguish this case from [A]? An at-first-sight attractive and well-trodden path would be to say that Mates's puzzle and its cousins involve a different semantic mechanism than Frege's puzzle, which generates our case [A] (see for instance Kripke, 1979, nn. 23 and 46). It is easy to construct versions of Frege's puzzle—confusion about an individual, a thing—for non-linguistic creatures. By contrast it is harder to construct plausible examples of Mates's puzzle (which seems to involve confusion about higher-type entities) for them. Citing this evidence, deep opacitists might follow a distinguished tradition which draws a line between Mates's puzzle—which may involve metalinguistic belief—and Frege's puzzle (for recent versions of this view, see Soria Ruiz, 2020, and Tancredi and Sharvit, 2020, with citations in the former). They might say that $Wa =_t Ga$, and that the reports which seem to distinguish these two are in fact based on some distinctive poor behavior of English attitude reports, which is not exhibited in the attitudes themselves.

This second style of approach has its own challenges. As I understand the position, it is in part motivated by the idea that non-trivial identities among *et*-properties play an important explanatory role (otherwise, why not hold that: is salt $\neq_{e \to t}$ is *NaCl*?). But if nontrivial identifications among *et*-properties play this role, shouldn't nontrivial identifications among *t*-propositions play a similar one? The problem is that, if they do, it seems to push us toward accepting [B], which deep opacitists reject. The challenge for this second version of deep opacitism is to explain why *et*-identifications are important if *t*-identifications are not. Perhaps they might claim that scientific explanations that appear to end with *t*-propositional identities can be understood as in fact appealing to neighboring *et*-property identities. But working out such a view is a non-trivial task.

Unlike wide opacitism, deep opacitism allows us to uphold logical principles governing $=_t$ and $\forall_t$. But there is a great deal of work to be done to understand which if any version of the position can be developed into an attractive overall view.

Acknowledgements

I would like to thank Kyle Blumberg, Ray Buchanan, David Builes, Tim Button, Mike Caie, Sam Carter, Josh Dever, Cian Dorr, Lidal Dror, Simon Goldstein, Jeremy Goodman, Ben Holguín, David Sosa, Una Stojnić, Rob Trueman, and especially Peter Fritz and Nick Jones for myriad helpful comments and conversations. Peter Fritz and Nick Jones have shown the patience of saints in awaiting this paper through a series of long delays.

Bibliography

Mahrad Almotahari and Aidan Gray. Frege cases and bad psychological laws. *Mind*, 130(520): 1253–1280, 2021.

Andrew Bacon. The broadest necessity. *Journal of Philosophical Logic*, 47(5):733–783, 2018.

Andrew Bacon. Logical combinatorialism. *Philosophical Review*, 129(4):537–589, 2020.

Andrew Bacon. Opacity and paradox. *Modes of Truth: The Unified Approach to Truth, Modality, and Paradox*. Routledge, 2021.

Andrew Bacon. A theory of structured propositions. *Philosophical Review*, 133(2): 173–238, 2023.

Andrew Bacon and Cian Dorr. Classicism, please add page range.

Andrew Bacon and Jeffrey Sanford Russell. The logic of opacity. *Philosophy and Phenomenological Research*, 99(1):81–114, 2019.

George Bealer. Property theory: The type-free approach v. the Church approach. *Journal of Philosophical Logic*, 23(2):139–171, 1994.

Francesco Berto and Daniel Nolan. Hyperintensionality. In *Stanford Encyclopedia of Philosophy*. 2021.

Kyle Blumberg and Harvey Lederman. Revisionist reporting. *Philosophical Studies*, 2020.

David Braun. Understanding belief reports. *The Philosophical Review*, 107(4):555–595, 1998.

David Braun. The objects of belief and credence. *Mind*, 125(498):469–497, 2016.

Michael Caie, Jeremy Goodman, and Harvey Lederman. Classical opacity. *Philosophy and Phenomenological Research*, 101(3):524–566, 2020.

Ben Caplan. Ontological superpluralism. *Philosophical Perspectives*, 25(1):79–114, 2011.

Rudolf Carnap. *Meaning and Necessity: A Study in Semantics and Modal Logic*. University of Chicago Press, 1947.

David J. Chalmers. Frege's puzzle and the objects of credence. *Mind*, 120(479):587–635, 2011.

David J. Chalmers. Referentialism and the objects of credence: A reply to Braun. *Mind*, 125(498):499–510, 2016.

Mark Crimmins and John Perry. The prince and the phone booth: Reporting puzzling beliefs. *The Journal of Philosophy*, 86(12):685–711, 1989.

Justin D'Ambrosio. Prior's puzzle generalized. *Philosophy and Phenomenological Research*, 106(1):196–220, 2023.

Donald Davidson. On saying that. *Synthese*, 19(1/2):130–146, 1968.

Cian Dorr. Transparency and the context-sensitivity of attitude reports. In Manuel Garcia-Carpintero and Genoveva Martí, editors, *Empty Representations: Reference and Non-Existence*, pages 25–66. Oxford University Press, 2014.

Cian Dorr. To be F is to be G. *Philosophical Perspectives*, 30:39–134, 2016.

Cian Dorr, John Hawthorne, and Juhani Yli-Vakkuri. *The Bounds of Possibility: Puzzles of Modal Variation*. Oxford University Press, 2021.

Hartry Field. Attributions of meaning and content. In *Truth and the Absence of Fact*, pages 157–174. Oxford University Press, 2001.

Hartry Field. Egocentric content. *Noûs*, 51(3):521–546, 2017.

Graeme Forbes. *Attitude Problems: An Essay on Linguistic Intensionality*. Clarendon Press, 2006.

Graeme Forbes. Content and theme in attitude ascriptions. In Alex Grzankowski and Michelle Montague, editors, *Non-Propositional Intentionality*, pages 114–133. Oxford University Press, 2018.

Gottlob Frege, Peter Geach, and Max Black. *Translations From the Philosophical Writings of Gottlob Frege*. Blackwell, 1970.

Peter Fritz. How fine-grained is reality? *Filosofisk Supplement*, 13(2):52–57, 2017.

Peter Fritz. Ground and grain. *Philosophy and Phenomenological Research*, 105(2): 299–330, 2022.

Peter Fritz. Operands and instances. *Review of Symbolic Logic*, 16(1): 188–209, 2023.

Peter Fritz. *The Foundations of Modality*, Oxford University Press, 2023.

Peter Fritz, Harvey Lederman, and Gabriel Uzquiano. Closed structure. *Journal of Philosophical Logic*, 50(6):1249–1291, 2021.

Jeremy Goodman. Reality is not structured. *Analysis*, 77(1):43–53, 2017.

Jeremy Goodman. Agglomerative algebras. *Journal of Philosophical Logic*, 48(4):631–648, 2018.

Jeremy Goodman. Grounding generalizations. *Journal of Philosophical Logic*, 52(3):821–858, 2023.

Jeremy Goodman and Harvey Lederman. Perspectivism. *Noûs*, 55(3):623–648, 2021.

Hüseyin Güngör. *That* solution to Prior's puzzle. *Philosophical Studies*, 179(9):2765–2785, 2022.

Richard Kimberly Heck. Solving Frege's puzzle. *Journal of Philosophy*, 109(1–2):132–174, 2012.

Nicholas K. Jones. Nominalist realism. *Noûs*, 52(4) 808–835, 2018.

Nicholas K. Jones. Propositions and cognitive relations. *Proceedings of the Aristotelian Society*, 119(2):157–178, 2019.

David Kaplan. *Foundations of Intensional Logic*. PhD thesis, UCLA, 1964.

David Kaplan. Quantifying in. *Synthese*, 19:178–214, 1968.

David Kaplan. Opacity. In Lewis Edwin Hahn and Paul Arthur Schilpp, editors, *The Philosophy of W. V. Quine*, pages 229–289. Open Court, 1986.

Kevin C. Klement. *Frege and the Logic of Sense and Reference*. Routledge, 2001.

Angelika Kratzer. Decomposing attitude verbs. *Talk given at the Hebrew University of Jerusalem*, 2006.

Saul A Kripke. A puzzle about belief. In Avishai Margalit, editor, *Meaning and use*, pages 239–283. D. Reidel, 1979.

Harvey Lederman. Fine-grained semantics for attitude reports. *Semantics and Pragmatics*, 14(1), 2021.

Harvey Lederman. Fregeanism, sententialism, and scope. *Linguistics and Philosophy*, 45(6):1235–1275, 2022.

Jon Erling Litland. A note on the wilhelmine inconsistency. *Analysis*, 81(4):639–647, 2022.

Brian Loar. Reference and propositional attitudes. *Philosophical Review*, 81(1):43–62, 1972.

Benson Mates. Synonymity. In Leonard Linsky, editor, *Semantics and the Philosophy of Language*, pages 111–136. University of Illinois Press, 1952.

Robert J Matthews. That-clauses: Some bad news for relationalism about the attitudes. *Mind & Language*, 37(3):414–431, 2020.

Kris McDaniel. *The Fragmentation of Being*. Oxford University Press, 2017.

Friederike Moltmann. Propositional attitudes without propositions. *Synthese*, 135(1): 77–118, 2003.

Keir Moulton. *Natural Selection and the Syntax of Clausal Complementation*. PhD thesis, University of Massachusetts, 2009.

Keir Moulton. CPs: Copies and compositionality. *Linguistic Inquiry*, 46(2):305–342, 2015.

Reinhard Muskens. Hyperfine-grained meanings in classical logic. *Logique et Analyse*, 133:159–176, 1991.

Reinhard Muskens. Sense and the computation of reference. *Linguistics and Philosophy*, 28(4):473–504, 2004.

Orin Percus and Uli Sauerland. On the LFs of attitude reports. In Matthias Weisgerber, editor, *Proceedings of Sinn und Bedeutung*, 2003.

Paul M. Pietroski. *Causing Actions*. Oxford University Press, 2000.

Paul M. Pietroski. Meaning before truth. In Gerhard Preyer and Georg Peter, editors, *Contextualism in Philosophy: Knowledge, Meaning, and Truth*. Oxford University Press, 2005.

A. N. Prior. Oratio obliqua. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 37:115–126, 1963.

Arthur N. Prior. *Objects of Thought*. Clarendon Press, 1971.

Willard V Quine. Quantifiers and propositional attitudes. *The Journal of Philosophy*, 53(5):177–187, 1956.

Willard Van Orman Quine. *Word and Object*. MIT Press, 1960.

Frank Plumpton Ramsey. *The Foundations of Mathematics and Other Logical Essays*. Routledge and Kegan Paul, 1931.

Agustín Rayo and Stephen Yablo. Nominalism through de-nominalization. *Noûs*, 35(1):74–92, 2001.

Tobias Rosefeldt. 'That'-clauses and non-nominal quantification. *Philosophical Studies*, 137(3):301–333, 2008.

Nathan U. Salmon. *Frege's Puzzle*. MIT Press, 1986.

Jennifer M Saul. *Simple Sentences, Substitution, and Intuitions*. Oxford University Press, 2010.

Stephen Schiffer. Naming and knowing. In A. French Peter, E. Uehling Theodore, Howard Jr, and K. Wettstein, editors, *Midwest Studies in Philosophy II: Contemporary Perspectives in the Philosophy of Language*, pages 28–41. University of Minnesota Press, 1979.

Scott Soames. Direct reference, propositional attitudes, and semantic content. *Philosophical Topics*, 15(1):47–87, 1987.

Andrés Soria Ruiz. On Mates' puzzle. Mind and Language 38(2):515–544, 2023.

Christopher Tancredi and Yael Sharvit. Qualities and translations. *Linguistics and Philosophy*, 43(3):303–343, 2020.

Richmond H. Thomason. A model theory for propositional attitudes. *Linguistics and Philosophy*, 4(1):47–70, 1980.

Robert Trueman. *Properties and Propositions: The Metaphysics of Higher-Order Logic*. Cambridge University Press, 2021.

Robert Trueman. Truth without dependence. *Aristotelian Society Supplementary Volume*, 96(1):89–121, 2022.

Jason Turner. Ontological pluralism. *Journal of Philosophy*, 107(1):5–34, 2010.

Gabriel Uzquiano. Ramified structure. *Philosophical Studies*, 80(5–6):1651–1674, 2022.

Peter van Elswyk. Propositional anaphors. *Philosophical Studies*, 176(4):1055–1075, 2019.

Peter van Elswyk. "That"-clauses and propositional anaphors. *Philosophical Studies*, 177(10):2861–2875, 2020.

Peter van Elswyk. The linguistic basis for propositions. In Chris Tillman, editor, *The Routledge Handbook of Propositions*. Routledge, 2022.

Isaac Wilhelm. Grounding and propositional identity. *Analysis*, 81(1):80–81, 2021.

Timothy Williamson. Everything. *Philosophical Perspectives*, 17:415–465, 2003.

Timothy Williamson. Epistemological consequences of Frege puzzles. *Philosophical Topics*, 49(2):287–319, 2021.

Timothy Williamson. Metametaphysics and semantics. *Metaphilosophy*, 53(2–3):162–175, 2021.

Seth Yalcin. Quantifying in from a Fregean perspective. *Philosophical Review*, 124(2): 207–253, 2015.

Juhani Yli-Vakkuri and John Hawthorne. Intensionalism and propositional attitudes. *Oxford Studies in Philosophy of Mind*, 2:114, 2021.