Synthese (2025) 205:193
https://doi.org/10.1007/s11229-025-05017-3

ORIGINAL RESEARCH

Frege's problem psychologized, concept atomism, and the division of labor

Kevan Edwards[1]

Received: 22 March 2024 / Accepted: 22 March 2025

Abstract
The main goal of this paper is to defend so-called *atomist* approaches to concept individuation against the threat presented by what I will refer to as *Frege's Problem Psychologized* (FPP): difficulties presented by putative cases of co-referring but distinct concepts. The discussion will provide an opportunity to highlight the virtues of a particularly austere, reference-based version of Concept Atomism and to draw attention to some broader morals, notably that even a radical version of atomism is consistent with embracing various features of cognition that are of interest to psychologists and empirically oriented philosophers.

Keywords Cognitive structures · Concepts · Concept atomism · Frege's problem · Reference · Representation

1 Introduction

I will follow a standard set of assumptions to the effect that *concepts* are psychological entities, the roughly word-sized constituents of propositional attitudes, and that philosophers and psychologists are engaged in a joint project, or at least working in parallel, to discover what concepts are, how they are individuated, and what theoretical role(s) they can and should play.[1] I will understand *Concept Atomism* as the family

[1] Laurence and Margolis (1999b) continues to be a standard point of entry into the topic of concepts; also see the more recent *Stanford Encyclopedia of Philosophy* entry on concepts by the same authors. A few details are worth acknowledging outside of the main body of the text. (1) By using the term "concept" to pick out a class of psychological entities, I am setting aside a question about whether concepts are best construed as mental representations or as some kind of abstracta (Laurence & Margolis, 2008), although I am not precluding the possibility that embracing the relevant mental representations brings a corre-

✉ Kevan Edwards
kedwar02@syr.edu

[1] Syracuse University, Syracuse, NY, USA

Published online: 22 April 2025

of theories about concept individuation that remains if one denies that the identities of so-called basic concepts stem from features of their internal structure and/or relations to other (conceptual or non-conceptual) features of cognition. This construal of Concept Atomism is both more specific, and in some ways more cautious, than some standard glosses in the literature, notably the claim that basic concepts lack internal structure.[2] A few details are worth clarifying. First, my construal of Concept Atomism does not preclude basic concepts from having structure; what it rules out is that structural features contribute to the identity of these concepts. Second, as I am construing the approach, Concept Atomism stands opposed not only to views that tie the identity of concepts to any internal structural features but also to what one might describe as their "external" structural features, namely their immediate relations to other concepts or to further (non-conceptual) features of cognition.[3] It goes beyond the scope of the present paper to defend the claim that this is a dialectically illuminating way to construe Concept Atomism; suffice it for now to suggest that this construal falls out of many of the most influential arguments that have been given in support of Concept Atomism.[4]

Despite Fodor's (e.g. 1998, 2008) fierce advocacy, Concept Atomism has remained at the periphery of philosophical work on concepts and has had little effect on discussions of concepts in cognitive and developmental psychology.[5] I submit that a cluster of related issues have played a significant role in fueling skepticism about Concept Atomism. At a very general level, Concept Atomism is deemed too far removed from issues that cognitive psychologists and empirically oriented philosophers take to be closely related to the notion of a concept.[6] Prinz (2002), for example, follows a standard assumption that making sense of *categorization*, as this notion is understood by psychologists, is a key desideratum for a theory of concepts, about which he says the following: "The greatest shortcoming of atomism involves categorization. Unstruc-

sponding class of abstract objects in its wake. (2) questions can be raised about whether philosophers and psychologists are operating with the same notion of a concept and even whether the term corresponds to a unified class of entities or, instead, if the notion should be eliminated. For discussion, see Machery (2009) and the various commentaries published alongside Machery (2010), as well as Piccinini and Scott (2006) and Weiskopf (2009).

[2] Thanks to an anonymous referee for drawing attention to the fact that I am taking some liberties with how I construe Concept Atomism.

[3] One reason to work with such a liberal notion of "structure" is that the line between features of a concept's role and claims about its structure is not always clear. For example, talk of concepts having prototypical structure, or even just *being prototypes*, might be explicated in terms of claims about its role, namely that the concept is deployed in a way that is driven by the activation of more basic, "feature" representations.

[4] The understanding of Concept Atomism that I am working with is similar in relevant respects to a suggestion made by Laurence and Margolis (1999a) after they clarify various notions of "structure" that the Concept Atomist might want to embrace. Also see Laurence and Margolis (2020), who gloss Concept Atomism in a way that explicitly runs together claims about a concept's structure with claims about its relations to other concepts.

[5] As noted by Kwong (2007), Concept Atomism is not even mentioned in the highly influential, *Big Book of Concepts* (Murphy, 2002).

[6] There are exceptions: Kwong (2007), Quilty-Dunn (2021), Brody and Feiman (2023) defend versions of Concept Atomism and are specifically interested in the relationship between atomism and empirically-oriented work.

tured mental representations simply cannot explain how we categorize" (98)[7]. Relatedly, there is a widespread assumption, one so deep that it is often not made explicit, that whatever concepts are, they must be some or another kind of cognitive *structure* (e.g. prototypes, exemplars, causal models, perceptual representations) or a combination or construction thereof.[8] Again, insofar as the atomist denies that concepts should be individuated in a way that appeals to such structural features, they might appear to be setting aside what others take to be the subject matter. By the end of the present paper, it will be clear why I think this is a mistake.

A closely related issue has to do with the austerity of the atomist's explanatory resources. Insofar as the atomist denies that concepts have either constitutive structure or constitutive relations to other concepts or to further aspects of cognition, the view might appear to be hobbled by limited explanatory resources. A way to make this concern vivid begins with the fact that a natural way to flesh out an atomist approach is to lean heavily on a notion of reference or representation.[9] This raises a host of concerns about the coarse-grained nature of reference that have been familiar at least since seminal work by Frege (e.g. 1948[1892]) and Russell (e.g. 2005[1905]). An obvious manifestation of this general concern is the worry that the Concept Atomist will run into the psychological analogue of Frege's Problem—what I will refer to as *Frege's Problem Psychologized (FPP)*. More specifically, it appears that the atomist will be unable to distinguish members of various pairs of putatively distinct but co-referring concepts: HESPERUS versus PHOSPHORUS, WATER versus H2O, TRIANGLE versus TRILATERAL, and so on.[10,11]

The main goal of this paper is to disarm the threat that FPP is assumed to present for atomist approaches to concepts.[12] My strategy will be to work within the confines of a version of Concept Atomism that is radical in its austerity. The *Concept Referentialist* claims that concepts are individuated by their contents and that conceptual content is exhausted by a notion of reference or representation. The Concept Referentialist thereby embraces a relatively coarse-grained notion of a concept, denying, for

[7] See Edwards (2009) for an explicit response to Prinz (2005) on this issue.

[8] This assumption is made explicit by Machery (2009).

[9] For the purposes of this paper, I am going to run together talk of "reference" and talk of "representation". I will clarify some working assumptions about the nature of reference later in the paper.

[10] Here and throughout, I follow a standard convention of using expressions in caps to indicate names of concepts.

[11] Another, arguably related, problem facing the Concept Atomist is how to handle so-called "empty" or non-referring concepts. See Edwards (2010) for a response that is consistent with the approach to FPP articulated in the present paper.

[12] A piece of anecdotal evidence about the polemical importance of FPP: virtually every conversation that I have had with another philosopher about the prospects of Concept Atomism has drifted onto the topic of how the atomist can hope to address Frege's Problem. There are a few notable defenses of atomism against (what I am labelling) FPP in the literature. While the present paper will diverge from Fodor's approach to FPP in important ways, it owes much to arguments that Fodor uses to motivate Concept Atomism as well as Fodor's (2008, ch. 3) strategy of dealing with putative hard cases by invoking a catalogue of resources—Fodor describes these as taking "nibbles" out of the problem. Rives (2009) discusses FPP and its relationship to Fodorian atomism, and in my view rightly emphasizes the importance of distinguishing between referential content and facts about the internal role of a concept. Rives ends up arguing for what he describes as a "compromise" between Fodorian "Cartesianism" and a form of pragmatism that I follow Fodor in rejecting.

example, that "HESPERUS" and "PHOSPHORUS" are names of distinct concepts.[13] In biting this bullet, the approach distinguishes itself from versions of atomism that end up qualifying the austerity of the approach in the face of FPP. Notably, even Fodor (2008, ch. 3) retreats in the face of FPP by appealing to resources such as *formal properties* and *mental files* in how he individuates concepts. I will return to Fodor's view in § 7, but suffice it to say for now that while the Concept Referentialist can embrace formal properties and architectural structures such as mental files, they deny that these features make a direct contribution to how concepts are type-individuated. Indeed, a core burden of the overall argumentation of this paper is to show how the Concept Atomist can appeal to various resources for the purposes of resolving FPP while denying that they play a direct role in the individuation of concepts.

One reason for working within the confines of such an austere theory of concepts is strategic: Concept Referentialism amounts to a worst-case scenario with respect to FPP and related issues. If the Concept Referentialist has access to resources sufficient to handle the relevant problem cases, it is reasonable to conclude that FPP does not present a serious problem for Concept Atomism more generally. A related point is that if the Concept Referentialist can deal with FPP, this undermines the need for atomists to walk back the austerity of their approach in the face of putative problem cases. This is a good thing, given the motivations for austerity that will be outlined in § 2 and the dialectical instability of less austere approaches that will be discussed in § 6.

Here is a map of the paper. § 2 sketches out some features of the broader dialectical geography that will play a role in the paper. These help to clarify Concept Atomism and to motivate the especially austere version of the approach endorsed in the present paper. § 3 and § 4 lay out a catalogue of resources that can be used to resolve various instances of FPP and that are not only consistent with Concept Referentialism but natural companions for the approach. § 5 addresses the issue of what notion of a *type of concept* a theory of concepts needs to vindicate. This yields a more sophisticated picture of the basic commitments behind Concept Atomism and addresses a potential concern to the effect that the resources invoked in the previous two sections undermine a commitment to Concept Atomism. § 6 highlights the *persistence* of FPP and uses this to motivate two features of the present approach that otherwise might raise suspicions. This feeds into the discussion in § 7 of how and why the version of atomism articulated in this paper differs from that of other, more permissive atomists, notably Fodor. § 8 draws some broader morals, ending with the clarification that even a very radical version of atomism is consistent with embracing features of cognition that are of interest to psychologists and empirically oriented philosophers.

[13] Although see the discussion in § 5 that clarifies talk of the same or distinct concepts by distinguishing the notion of a *concept-type* from the more generic notion of a *type of concept*.

2 Concept referentialism versus role/structure theories

Given sufficiently broad notions of a concept's *role* and *structure*, most approaches to concepts, in both philosophy and psychology, can be lumped under the heading of what I will call *Role/Structure theories*: approaches that individuate concepts in virtue of features of their roles or structure (broadly construed). This broad heading includes so-called "classical" or "definitional" approaches, as well as currently popular approaches that identify concepts with various kinds of cognitive structures: prototypes, stereotypes, exemplars, as well as pluralist and hybrid constructions thereof.[14] Role/Structure theories also include inferential role theories of concepts (e.g. Peacocke, 1992), approaches that appeal to a concept's role in a theory (e.g. Carey, 2009), and empiricist approaches that appeal to a concept's role in or relation to perceptual systems or perceptual representations (e.g. Barsalou, 1999; Prinz, 2002). In short, "Role/Structure theory" is a broad heading that encompasses more or less all approaches to concepts aside from Concept Atomism.

An important clarification is that Role/Structure theories take the relevant features of concepts to be directly implicated in the type-individuation of concepts, or in other words to be *constitutive* features of concepts.[15] A recurring theme of the present paper is that while the Concept Atomist denies that various role/structure features play a direct role in concept identity, they need not deny that concepts have such features or even that they play what amounts to an indirect role in concept individuation. To gesture towards a possibility that will be fleshed out in § 4, the Concept Referentialist might claim that the role/structure features of concepts help to realize the relation that they take to be concept-constitutive, namely reference.

I assume there is a workable distinction between facts about a concept's role/structure (again, broadly construed) and its atomistic properties, notably its referential content.[16] I thus assume that an atomist-*cum*-referentialist approach to concepts is a clear alternative to Role/Structure theories.[17] This means that general arguments against concept-constitutive appeals to role/structure features can be counted in favor

[14] Prototype approaches to concepts (e.g. Hampton, 2006) are widely acknowledged to have their roots in seminal work by Rosch (e.g. 1975), although Rosch does not make a claim about concepts per se. Smith and Medin (1981) is a now classic introduction to psychological theories of concepts that clarifies different approaches that can be lumped under the heading of "prototype" and "exemplar" theories. For a recent defense of a so-called *hybrid* approach, see Vicente & Martinez Manrique (2016). For a defense of *concept pluralism*, see Weiskopf (2009).

[15] I follow a use of "constitutive" that is familiar in the literature on concepts. I make no claim about it being consonant with other uses of the term and I do not know of any helpful discussions of this issue.

[16] It is possible to understand the notion of conceptual role so broadly that it encompasses claims about reference (e.g., Harman & Greenberg, 2006). Insofar as the issues here are terminological, I submit that such a permissive construal of conceptual role is not a good choice insofar as it obscures important and contested theoretical issues.

[17] A few issues are worth noting here. One is a reminder that I have deferred a discussion of formal properties and mental files to § 7. Another is that while Concept Atomism is at odds with Role/Structure theories of *basic* concepts, Concept Atomism is naturally combined with the claim that *non-basic* concepts are compositional constructions, a point that will be crucial to the discussion in § 3. This raises the question of how to characterize the distinction between basic and non-basic concepts. One might try to finesse this by appealing to "lexical" concepts, but of course this is just another label that ultimately needs to be explicated.

of atomism, and in particular a reference-based version of atomism. Indeed, the arguments that have done the most to motivate Concept Atomism are due to Jerry Fodor and exploit this feature of the dialectical geography.

One way to sum up Fodor's arguments in support of Concept Atomism is that they accuse alternative approaches (what I am lumping under the heading of "Role/Structure theories") of violating various "non-negotiable" (Fodor, 1998) constraints on theories of concepts, notably *shareability* and *compositionality*.[18] In crude terms—it will help to have a placeholder going forward—what makes it difficult for non-atomistic approaches to satisfy shareability and compositionality constraints on concepts is the *messiness* of role/structure features.[19] Facts about the structure and role of a particular concept can be sensitive to issues that are potentially complicated, wide-ranging, and flexible: facts about its possessors' background beliefs and preferences, the nature of the specific processes and/or structures in which the concept is embedded, and so on. To make a familiar Quinean point, an otherwise rational person may be willing to draw prima facie crazy inferences involving a particular concept against the right background of (false) beliefs. This complexity, plasticity, and situation-specificity of a concept's role and structure—again, what I want to capture by describing these features as *messy*—make it difficult to pin down features that are shared in a sufficiently robust way across agents, times, cognitive contexts, environmental conditions, etc. Similarly, insofar as a concept's role/structure features can and do change when it occurs as a constituent of a complex thought, Role/Structure theories have trouble vindicating a strong compositionality constraint.

In contrast with the messiness of role/structure features, the austerity of an atomist-*cum*-referentialist approach leads to a much more robust notion of a concept; this makes it relatively easy to vindicate shareability and compositionality constraints. As a first approximation, agents share a concept insofar as they possess concepts that refer to the same individual/kind/property. Working out the full details of the compositional structure of thought promises to be difficult for everyone; but, again, I take it to be relatively uncontroversial that the atomist-*cum*-referentialist has a comparatively easy time vindicating a compositionality constraint.

To sum up, it is arguably the contrast between the *messiness* of role-structure features of concepts and the relative *austerity* of an atomistic feature like reference that is at the heart of what are arguably the most influential arguments in favor of Concept Atomism and that motivate cleaving to a reference-based version of the approach.

I suggested in § 1 that instances of FPP are reasonably viewed as manifestations of more general concerns about the coarse-grained nature of atomistic approaches. Against this backdrop, worries about the messiness of role/structure theories adds a dialectically-interesting wrinkle. While instances of FPP might seem like a reason-

[18] Georges Rey has convinced me that "shareability" is a better term than the more traditional "publicity", notably because it avoids potentially misleading epistemic associations. The starting point for (what I am calling) the shareability and compositionality arguments is Fodor (1998). For an extended discussion of compositionality in the context of theories of concepts, see Fodor and Lepore (2002). Another argument that Fodor (e.g. 2004) wields against various approaches (what he refers to as "pragmatist" theories of concepts) involves enforcing a non-circularity constraint. This argument that will not play a role in the present paper.

[19] Quilty-Dunn (2021) uses the term "messy" in a similar way.

able way to motivate more fine-grained approaches to concept individuation, this promises to exacerbate worries about the messiness of those features.[20] I think this tension is instructive. The approach articulated in the present paper takes this tension seriously by cleaving to an especially austere approach to concepts but rejecting the assumption that FPP should be resolved by a theory of concepts per se.

A quick but important clarification about notation: I am following a standard practice of generating names of concepts by putting familiar expressions in caps. But this seemingly innocent move can make it easy to beg the question against the atomist. For example, insofar as the Concept Referentialist denies that pairs of expressions such as "CILANTRO" and "CORIANDER" pick out distinct types of concepts, they may—and I think they should—worry that the use of these expressions loads the deck against their approach. When it is helpful to capture the situation from the perspective of someone who denies the type-distinctness of a putative pair of concepts, I will use a slash notation. For example, in place of "CORIANDER" and "CILANTRO", I will run the discussion in terms of tokens of the type "CIL/COR".[21]

3 Three basic resources

The goal of this section is to briefly catalogue some relatively basic resources that are consistent with Concept Referentialism and that can be brought to bear on instances of FPP. I will suggest that while these resources promise to play a role in resolving various instances of FPP, they are not enough to count as a full solution. The emphasis will be on the availability and limits of these resources rather than staking firm claims about specific cases.

3.1 Minimal-LOT

[22]It is natural for the Concept Atomist to adopt at least a minimal version of the so-called Language of Thought Hypothesis.[23] What I will call *Minimal-LOT* is a claim about the semantically-relevant structure of thought, namely that it is compositional and that it roughly approximates the structure of familiar natural languages.[24] Minimal-LOT can do explanatory work for the Concept Referentialist. Consider a situ-

[20] Onofri (2016) argues that it is impossible for a theory of concepts to both underwrite a solution to (what I am calling) FPP and vindicate shareability.

[21] This apparently heavy use of the type/token distinction will be partially discharged in § 5.

[22] This brief discussion of what I am referring to as "Minimal-LOT" parallels a discussion in Fodor (2008, ch. 3), which goes into greater detail about the constituent structure assumed by the Language of Thought and how this can be brought to bear on various instances of FPP.

[23] There is a lot to be said both for and against the LOT hypothesis. For an extensive recent discussion, see Quilty-Dunn et al. (2023), as well as the accompanying commentaries and responses.

[24] "Minimal-LOT" is an attempt to exercise caution. The advocate of Minimal-LOT is not committed to the brain being populated by linguistic inscriptions, nor to natural language having any substantial (ontological, developmental, evolutionary) priority over thought, nor to any particular (e.g. computational) account of mental processes (*pace* what is assumed by Schneider, 2011). Even the claim that the structure of thought approximates that of natural language should be treated as a defeasible working assumption.

ation in which it is natural to describe someone as believing *that the 45th President of the United States is well informed* (after all, presidents have access to presidential briefings) even though it is natural to refrain from describing that same person as believing *that Donald Trump is well informed*. Plausibly, any reasonable way of fleshing out Minimal-LOT will involve enough structure that the Concept Referentialist will be able to distinguish the thoughts in question.

What is the scope of appeals to Minimal-LOT when it comes to resolving FPP? It seems safe to assume that an appeal to relevant conceptual structure can help to explain cases that involve not only descriptively-structured thoughts but just complex concepts, perhaps including TRI-ANGLE versus TRI-LATERAL and MORNING-STAR versus EVENING-STAR. But even here, controversy looms: the relationship between lexical items in natural language and corresponding concepts is far from clear (does "triangle" express TRI-ANGLE?), and there are ongoing debates about which concepts have internal structure.[25] Minimal-LOT has the potential to be a useful resource for responding to at least some instances of FPP, but it is worth exercising caution about how broadly it can be applied.

3.2 The gap between language and thought

Since putative problem cases are presented in natural language and often use distinctions in natural language to suggest related distinctions in thought, the relationship between expressions in natural language and corresponding conceptual resources has the potential to help resolve instances of FPP. Roughly speaking, there is room for various kinds of *gaps* between what is apparent in the surface structure of sentences that play a role in setting up a putative problem case and the underlying thoughts involved. Specific examples of "gaps" between language and thought are apt to engender controversy, but here are some familiar sentences paired with thoughts that one might claim the sentences are used to express/communicate in easy-to-imagine situations:[26]

(S3) Jack and Jill went up the hill.

(T3) JACK AND JILL WENT UP THE HILL TOGETHER

(S4) Stephen is ready.

[25] Relatively recent philosophical defenses of the claim that typical lexical concepts have significant structure include Daley (2010) and Vicente (2010); although, as Daley's discussion makes clear, care needs to be taken with the relevant notion of *structure*. Count the present author among those who doubt that typical lexical items in natural language have definitions, never mind definitions in the Language of Thought. For a seminal contribution to this side of the debate, see Fodor et al. (1980); also see Fodor's (1981b) infamous failure to define the transitive verb "to paint".

[26] I assume that the standard convention of putting names of concepts in caps can be extended to whole thoughts in a way that makes clear some working assumptions about the underlying semantically-relevant structure of those thoughts. S3 is taken from Bach (2001), S4 is from Cappelen and Lepore (2005), S5 is from Perry and Blackburn (1986). I do not mean to suggest that any of these authors would endorse what I am (tentatively) suggesting about the structure of the thoughts in question.

(T4) STEPHEN IS READY FOR THE PARTY

(S5) It is raining.

(T5) IT IS RAINING IN MY CURRENT LOCATION

For present purposes, I shy away from claims about the specific mechanisms underlying these or other such "gaps". As far as resolving instances of FPP is concerned, I am inclined to doubt that much explanatory work can be squeezed out of what one might call *semantic gaps*: roughly speaking, differences between what an expression might appear to mean on the surface and the concepts/thoughts that it *semantically expresses* or that encode its *literal meaning*. I am not optimistic, for example, that standard cases involving "Hesperus" and "Phosphorus" can be explained by claiming that such expressions inherit their semantic content from complex concepts such as EVENING-STAR and MORNING-STAR.

A safer bet is that gaps resulting from something looser than a semantic relation between language and thought—call these *pragmatic gaps*—can be brought to bear on various (putative) instances of FPP. I assume that speakers in appropriate conversational contexts can and do get away with using proper names to communicate descriptively rich thoughts, even if this outstrips the literal semantic content of the expressions so used. This can lead to situations that might appear on the face of their linguistic descriptions to be a problem for the Concept Atomist but in fact involve thoughts that even the Concept Referentialist can distinguish. Take a familiar kind of case involving Byron-the-Bystander who, witnessing the way that Lois Lane interacts with Superman/Clark Kent in various kinds of situations, assertively utters each of the following sentences:[27]

(S6) Lois Lane believes that Superman can fly.

(S7) Lois Lane believes that Clark Kent cannot fly.

This has the trappings of an instance of FPP; it might appear that making sense of Byron's reason for uttering these sentences will require attributing to Byron thoughts that involve distinct but co-referring concepts: SUPERMAN and CLARK-KENT. But perhaps this is too quick. In the right kind of context, Byron's willingness to utter these sentences might be explained by his intending the sentences, as uttered in the relevant context, to convey something like the following (descriptively enriched) thoughts, thoughts that the Concept Referentialist can hope to distinguish:[28]

(T6) LOIS LANE BELIEVES THAT SUP/CLARK-KENT—THE SUITABLY-ENCOSTUMED-SUPERHERO—CAN FLY

[27] I follow an unstated convention in much of the literature and assume it is possible to abstract away from issues stemming from the non-existence of Superman/Clark Kent.

[28] I am finessing various potentially controversial issues regarding the content of belief reports, both having to do with the thoughts expressed by belief reports and the thoughts thereby attributed.

(T7) LOIS LANE BELIEVES THAT SUP/CLARK-KENT—THE MILQUE-TOAST REPORTER—CANNOT FLY

It is worth pointing out something that will be discussed further in § 6, namely that the atomist need not have a uniform response to various cases of FPP, or even to various cases involving "Superman" and "Clark Kent". This is especially salient in the context of appeals to gaps between language and thought that are due to potentially messy pragmatic mechanisms.

It might be misleading to describe intentionalist gaps as themselves providing a resolution to various instances of FPP. In many cases, what they promise is to extend the scope of other resources. In the above case, a gap between the literal meaning of various belief ascriptions and what a speaker is intending to use the sentences to communicate makes room for an appeal to distinct conceptual structures—descriptively enriched thoughts.

3.3 The intensionality of reference

It will ease exposition, in this section and later in the paper, to work with a cartoon view about the metaphysics of reference. The approach that I will provisionally adopt is in keeping with what are sometimes called *informational* approaches to mental representation and are often allied with atomism.[29] I will assume that reference is a species of *reliable covariation*, where the relevant notion of covariance is law-like and hence can be illuminated by considering relevant counterfactuals. To a first approximation, a concept refers to the individual/kind/property that does/would cause it to be activated in a range of actual and possible situations. Using some familiar jargon, a concept refers to the individual/kind/property that causes it to be activated in a restricted set of possible worlds.[30]

A potential feature of the reference relation that is easy to illustrate using this approach is that concepts can refer to kinds or properties whose extensions coincide in the actual world but come apart in relevant possible worlds. Consider the conceptual analogue of Quine's (1970) famous case: the concepts RENATE and CORDATE. Following tradition, let's assume that every actual renate (creature with a kidney) is a cordate (creature with a heart), and vice versa. Nevertheless, as long as the metaphysics of reference is sensitive to the *possibility* of renates that are not cordates (and vice-versa), the concepts in question can have distinct referential contents and even the Concept Referentialist can recognize them as distinct.

There are alternative approaches to the metaphysics of reference that are in the same broad family of the nomic covariationist approach that I just invoked, for example attempts to ground reference in the *proper function* (of so-called *producers* or *consumers*) of a representation, where the notion of a proper function is analyzed along evolutionary lines (e.g. Millikan, 1989; Neander, 2017). For present purposes, I suggest that however a theory of reference/representation is developed, there is likely

[29] This is the case with both Fodor (e.g. 1992b) and Millikan (e.g. 2000).

[30] Yes, there are familiar and much-discussed problems with this kind of approach, not the least of which is cashing out the relevant notion of a "range" or "restricted set" of worlds in a non-circular way.

to be pressure to construe reference as an intensional relation and that this has the potential to be a tool that the atomist-*cum*-referentialist can bring to bear on at least some putative instances of FPP.

Thus far I have introduced three basic resources that can be brought to bear on putative instances of FPP. I do not think these are enough to count as a complete solution. Notice that standard problem cases turn on the intuitive distinctness of putative concept types that are plausibly unstructured and necessarily co-extensive: HESPERUS versus PHOSPHORUS, CILANTRO versus CORIANDER, and so on. Appealing to "gaps" between language and thought also looks to be limited in scope; for one, what is arguably at issue when it comes to many putative instances of FPP is not something that can be unpacked in terms of how we describe or characterize someone's thoughts, but a distinction in the thoughts themselves. Thankfully, I have not yet touched on a resource to which the Concept Referentialist can appeal and that promises to be more general and more powerful than those discussed thus far.

4 Meaning and mental structures/processes

Start with an abstract point: one can coherently embrace a taxonomy of types of entities where the types abstract away from various features of those entities that nevertheless play a role in drawing explanatorily relevant distinctions between tokens of those types. Here is how this applies to the issues at hand: there is a position in logical space according to which role/structure features do not make a constitutive contribution to the type-individuation of concepts per se but nevertheless help to explain distinctions between various tokens of those types in a way that explains putative instances of FPP. I will flesh out this abstract possibility in two ways: First, I will take steps to underscore the extent to which (non-concept-constitutive) appeals to role/structure features have a natural home in a broader package of commitments that the Concept Referentialist is apt to take seriously. Second, if differences in role/structure features are to help resolve instances of FPP, such features should be poised to explain relevant features of the cases. That this is so will be illustrated later in this section, by considering a few intuitive cases.

Once again, it will help to be able to work with a cartoon picture, this time regarding how concepts fit into larger cognitive structures and systems. The upshot, in broad outline, is that it is natural for the Concept Referentialist to embrace a claim about the relationship between a concept's content and its role in mental structures and processes to the effect that aspects of the latter are reliably *in sync* with the former, but that this leaves room for partial synchronization failures that can then be used to explain various instances of FPP.

Let's assume a crude, classical framework according to which various cognitive systems process representations, including concepts, in a way that is driven by the so-called "local" properties of those representations.[31] Within this framework, it is natural to think of it being a core function of so-called *input systems* (e.g., visual, auditory, and other perceptual systems) to react to impinging external stimuli by activating

[31] This is a paraphrase of Fodor's (1981a) "Formality Condition".

relevant concepts, assigning them relevant local properties, and making the concepts available to "downstream" cognitive systems.[32] To be a bit more precise, it is part of the function of input systems to ensure that concepts have local properties such that downstream cognitive systems process these concepts in a way that is appropriate given the relevant perceptual stimuli—the individuals/kinds/properties to which the concepts refer. In other words, well-functioning input systems help to ensure that the roles of various concepts in cognitive structures and processes are *reliably in sync* with their semantic contents. The result is that while the Concept Referentialist draws a distinction between the properties in virtue of which concepts are type-individuated (i.e. reference) versus those that are relevant to cognitive structures and processes, they can nevertheless embrace the claim that cognitive structures and processes are *indirectly* sensitive to a concept's (referential) content.

Similar comments apply to cognitive structures and processes other than input systems. Mechanisms that are more directly responsible for behavior can be expected to function in a way that allows a cognitive agent to successfully navigate what it represents as being in its environment. Analogous points can be made about cognitive machinery that is not so obviously at the input/output interface, for example inferential mechanisms that generate new thoughts in a way that reliably preserves semantic properties, for example reference and truth.

This sense in which concepts fit into larger thoughts, cognitive structures, and mental processes in a way that is reliably in sync with their referential contents is not much more than a common-sense idea that falls out of the cartoon framework that I have been assuming. Can we do more to develop this notion of role-content synchronization? I think we should have modest expectations. The aforementioned *messiness* (§ 2) of facts about a concept's role/structure is apt to frustrate attempts to say something that is both illuminating and very general. Arguably, what is needed to go beyond a rough-and-ready notion of role-content synchronization is nothing short of a more carefully elaborated theory of reference along with a richer and more empirically-informed understanding of the relevant cognitive systems, structures, and processes. This noted, the difficulty of cleaning up a rough-and-ready notion of role-content synchronization does not present a serious problem as far as the work that I am about to ask role/structure features to do vis-à-vis instances of FPP. For example, insofar as various role/structure features of a concept depend on a potentially intractably rich network of background conditions, relevant claims about such features can (explicitly or tacitly) assume such conditions. In the absence of full information about these background conditions, claims about specific role/structure features will be speculative and defeasible—arguably highly so. But what matters as far as undermining the force of FPP is the plausible existence and relevance of such features, not that we can pin them down with certainty, nor that what applies to specific cases generalizes to others in a robust way.

[32] For simplicity, I ignore such complications as the likelihood of intervening layers of (non-conceptual) representations and various mental processes being defined over larger units (e.g. thought-sized entities). Moreover, talk of input systems *activating* token concepts and *assigning* them local properties is a placeholder; for example, a picture according to which input systems *select* concepts according to the relevant properties would do just as well for present purposes.

There is a further way in which role/structure features have a natural home in the overall approach that I have assumed. Return to the crude picture of the reference relation to which I gestured in § 3. According to the approach that I provisionally adopted, the reference relation is a relatively abstract, high-level (law-like) covariance relation. As such, it is a relation that will need to be *implemented* by relatively low-level, nuts-and-bolts mechanisms.[33] The key point is parallel to the picture painted above regarding role-content synchronization: once one considers plausible candidates for such mechanisms, various facts about the role and structural features of concepts—both intuitive and those requiring empirical discovery—emerge very naturally. Consider mechanisms that might be part of a suitably complex cognitive system and that plausibly help to ensure that tokens of the type WATER (to stick with the familiar example) reliably co-vary with water. So-called input systems should be expected to play this role: tokens of the concept WATER covary with water (the stuff) in part because input systems respond to the presence of water by outputting token concepts with local properties that are appropriate given that content. Arguably, it isn't just input systems that play a role in implementing the reference relation. Ensuring covariance plausibly involves non-sensory, theory-driven inferences; complex dispositions involving abilities to recognize and defer to experts; and so on. To make a long story short, the fact that concepts have role/structure features and fit into various larger cognitive structures and systems is what the Concept Referentialist should expect given what it will take, as a matter of empirical fact, to implement the reference relation.

What will turn out to be relevant to resolving FPP is not role-content synchronization per se, but cases where role-content synchronization breaks down. There are some obvious reasons why partial breakdowns are to be expected despite general pressure to preserve synchronization. An obvious source is just the fallibility of the relevant systems and processes, taken either separately or in terms of how they interact with one another or with features of the broader environment. To lump this under a broader heading, there are numerous sources of what we can describe as relevant *epistemic deficiencies*. Consider, for a few quick examples, false beliefs that arise as the result of perceptual systems performing in sub-optimal conditions, the use of efficient-but-fallible inferential heuristics, or just bad information coming in from one or another otherwise reliable source. It won't play a significant role in the cases to be discussed, but it is worth noting a different kind of cause for role-content synchronization failures, namely that concepts can be used in ways that are *intended* to fail to reflect their contents, for example in hypothetical or counterfactual reasoning.

The bearing of all of this on FPP can now be made more explicit. While the Concept Referentialist should maintain that reliable role-content synchronization is (in some sense) the norm, what we have just seen is that there is room for various tokens of a concept type to have features of their roles that are out of sync with their content. Of particular interest are cases where role/structure features of co-referring concepts diverge such that, at least in relevant respects, cognitive systems effectively treat the concepts as if they have different contents. In short, Concept Referentialist is con-

[33] This is the notion of a "sustaining mechanism" from Laurence and Margolis (1999a).

sistent with co-referring concepts being treated by cognitive systems as if they were distinct, the hallmark of a putative instance of FPP.

Consider a very familiar kind of case. At certain times of the year, Astra-the-Astronomer sees a bright object in the morning sky and thinks to herself that it is a planet; at other times of the year, she sees a bright object in the evening sky and thinks to herself that it is not a planet. Unbeknownst to Astra, there is only one object that she is thinking about: the planet Venus. Arguably, the Concept Referentialist is committed to Astra believing both (T8) and (T9).

(T8) HES/PHO IS A PLANET

(T9) HES/PHO IS NOT A PLANET

The Concept Referentialist needs to be able to explain how Astra can believe each of these contradictory thoughts. To begin with, it is baked into the case that Astra is epistemically deficient in relevant respects; after all, it is uncontroversial that she believes various thoughts that are false, for example *that a certain bright object in the morning sky is not a planet*. It is a plausible upshot of this, and other relevant epistemic deficiencies, that various tokens of the type HES/PHO have role/structure features that are out of sync with their (referential) content, to the point where some instances of HES/PHO end up being constituents of false beliefs. Indeed, the initial description of the case is enough to suggest plausible sketches of distinct inferential pathways such that Astra ends up tokening instances of both (T8) and (T9).[34] It is worth noting that similar resources can be used to explain how Astra fails to believe some tokens of what the Concept Referentialist classifies as the same thought (e.g. (T8)). This helps to make sense of why we are not only willing to attribute inconsistent beliefs to Astra, but to describe her as *failing to believe* that Hesperus is a planet.[35] An obvious difference between various tokens of the type (T8), for example, is that some are such that Astra would not be willing to believe them as a result of a particular pattern of inference that nevertheless plays a role in her believing others.[36]

A way to sum up relevant features of Astra's situation is that various tokens of the type HES/PHO in her psychology have roles *as if* they refer to distinct heavenly

[34] Notice that it is not a problem for the Concept Referentialist if these inferential pathways are themselves complicated. To repeat a recurring theme, the Concept Referentialist needs to be able to gesture to relevant differences; they do not need to be able to clean up these resources into robust identity conditions for concepts.

[35] Since Kripke (1979), it has been standard to distinguish cases involving seemingly inconsistent beliefs (Paderewski believes that Paris is beautiful and also believes that Paris is not beautiful) and those involving seemingly incompatible belief reports (Paderewski believes that Paris is beautiful but also does not believe that Paris is beautiful). In broad outline, the present approach to FPP deals with the former kind of case by explaining the putative inconsistency and it deals with the latter by explaining the claims of non-belief in terms of the agent not believing various tokens of the relevant type of thought.

[36] Here I have slipped into appealing to a counterfactual claim about the thoughts in questions. I think it is independently plausible that these kinds of appeals will be available. I submit that if the Role/Structure theorist is going to have any hope of individuating concepts at a sufficiently fine-grain to resolve FPP, she will need to be able to appeal to features of a concept's role in counterfactual situations. The Concept Referentialist should feel free to do likewise.

bodies. Astra's being in this situation is both consistent with Concept Referentialism and something that the Concept Referentialist can hope to explain, notably in terms of various uncontroversial epistemic deficiencies and the effects these have on the roles of various concepts that are nevertheless concepts of the same type (HES/PHO).

What happens when Astra gets further information such that it is natural to describe her as coming to realize that Hesperus is Phosphorus? According to the Concept Referentialist, this is not a matter of Astra coming to believe a thought of the form, HESPERUS IS PHOPHORUS, since this (putative) thought involves types of concepts that the Concept Referentialist rejects. In broad outline, the Concept Referentialist can explain this kind of case by maintaining that new information changes various facts about Astra's epistemic situation and that this is reflected in changes in the role/structure features of relevant concepts.[37] More specifically, various tokens of the type HES/PHO that her cognitive systems previously would have treated as if they had distinct contents would now be treated as having the same contents—at least as far as relevant issues are concerned.[38] For example, among other things that can be said about Astra's situation, there are tokens of the type HES/PHO IS HES/PHO that Astra's cognitive systems would previously have blocked from having the status of beliefs that are now free to play this role in her cognitive economy.

Consider another familiar kind of case: Bill-the-Believer is in the herbs & spices section of a grocery store and finds a small bottle labeled "Cilantro" and another labeled "Coriander". He has no reason to piece together that the contents of the bottles are from the same plant. Once again, a key question is whether the Concept Referentialist can distinguish various occurrences of what we can call "CIL/COR" and do so in a way that makes sense of how there can be various tokens of the same types of thoughts such that Bill is disposed to believe some, disbelieve others, and fail to believe yet others. Also once again, it is easy to gesture towards potential differences in the roles of various occurrences of CIL/COR that are plausibly relevant to explaining these facts. For example, consider various tokens of the thought CIL/COR IS CIL/COR that Bill would consider to be trivially true—those that one might characterize in practice using "Cilantro is Cilantro" or "Coriander is Coriander". As a first approximation, the tokens of CIL/COR that are constituents of these thoughts both co-refer and have convergent roles, at least when it comes to role/structure features that have the potential to make a difference to whatever issues are at hand. In other words, Bill's mental processes treat the concepts as if they co-refer, as indeed they do. In contrast, there will be other thoughts of the type CIL/COR IS CIL/COR that Bill is *not* willing to entertain as beliefs, namely those involving tokens of CIL/COR that have sufficiently divergent role features—roles as if they refer to different types of plants. As in the previous case, Bill is epistemically deficient in ways that the Concept Referentialist should expect to have this as a consequence. It is natural to describe Bill as failing to realize *that the contents of each of two jars he has picked*

[37] The Concept Referentialist can hope to go well beyond this broad outline, for example by invoking architectural features such as mental files. The availability of these kinds of notions will be clarified in § 7, but for now I set them aside for fear of fueling the concern that the Concept Referentialist is unfairly trying to have their cake and eat it too.

[38] To gesture to an issue to be discussed in the next section, Astra might still worry about further issues, for example that several different heavenly bodies might be catching her eye on different evenings.

up are from the same type of plant, and as assuming *that the words "Cilantro" and "Coriander" pick out different plants*, and so on. In general, it is natural to describe Bill as being unaware that what he takes to be two different plant species are in fact one. As in the previous case, the Concept Referentialist has obvious resources with which to explain Bill's cognitive situation vis-à-vis the various thoughts in question.

Appealing to differences in relevant features of a concept's role/structure, including differences that can arise in the case of co-referring concepts, looks to be a very general and very flexible resource, one that the Concept Referentialist can hope to bring to bear on various instances of FPP. Indeed, consider the nature of the problem cases that arguably slip through the cracks of the resources outlined in the previous section. Roughly speaking, these are cases where it is natural to think that an agent is employing concepts that, despite being unstructured and necessarily co-referring, are treated by the agent's cognitive systems as if they are distinct. Precisely what this section has focused on being able to explain is how co-referring concepts can end up playing a role in cognitive structures and processes as if they are distinct.

5 Types of concepts and concept-types

For the purposes of discussion, I am willing to accept that appealing to a property (or set of properties) for the purposes of some or another explanatory project is tantamount to recognizing a type defined by that property (or set thereof). Likewise, I accept that appealing to role/structure features of concepts in the process of distinguishing token concepts from one another amounts to acknowledging types of concepts that are individuated by those properties. Given these concessions, one might object that insofar as I have appealed to various role/structure features for the purposes of distinguishing various token concepts, I have effectively embraced a taxonomy of concepts that are type-individuated by facts about their roles/structures. And isn't this just to give up on Concept Atomism?

With respect to the accusation that I have invoked resources that can be used to define types of concepts that are individuated in terms of relevant role/structure features, I plead guilty with explanation. The important point is that I deny that this amounts to embracing a Role/Structure theory or giving up Concept Atomism. A related point is that running relevant discussions in terms of explicable distinctions between various concept tokens and failing to pay lip-service to the corresponding types of concepts has simplified the exposition, but it is not strictly speaking necessary. I now want to back up and be more explicit about these issues.

I have been taking for granted something that is widely, if often only tacitly, assumed in the literature on concepts: that the hunt for a theory of concepts is an attempt to capture some particular taxonomy of types of concepts, one that is privileged by those types playing an important explanatory role (or, more likely, roles).[39] I think this is a case where a natural way of speaking elides a theoretically significant issue. In everyday English, substitutions for "ϕ" in locutions of the form "types of

[39] Although see some of the details acknowledged in note 1.

ϕs"[40] often leave open various possible restrictions on the types at issue. To borrow an example from Block (1993, p. 57), we can type bathtubs by size, shape, weight, decorative properties, price, durability, insulating properties, expected life-span, thermal properties, etc. I suspect that in many situations we tacitly assume that resolving the type (of type) at issue is not likely to make a difference and we tolerate the indeterminacy. In situations where we intend to privilege one taxonomy of types over others, we either rely on this being obvious in the context or we make it explicit. In the present context, there is good reason to be careful about what type (of types) of concept we have in mind, precisely because there are various issues of theoretical interest on the line. If we aren't careful about the implications of different ways of typing concepts, the fact that there are various, potentially competing, explanatory pressures on a theory of concepts under discussion (notably: shareability, compositionality, and resolving FPP), can lead to cross-talk and confusion. In this context, then, it is arguably dangerous to run the discussion in terms of open-ended talk of "types of concepts" or even "concept individuation".

We can stipulate our way out of this problem. I reserve the term, "concept-type", to pick out the proper target of a theory of concepts per se, which gives us a way to distinguish this from other ways of typing concepts. Thus, "theories of concepts" should be understood as "theories of *concept-types*" and claims about how concepts are individuated should be understood as claims about how *concept-types* are individuated.

We can now be more careful about what the Concept Atomist is claiming and what distinguishes them from Role/Structure theorists. The Concept Atomist denies that concept-types are individuated in terms of the role/structure features of concepts, but they can happily embrace claims to the effect that concepts have such features, that concepts fit into broader cognitive structures and processes, and even that there are types of concepts that are individuated in virtue of such features. The dialectically important question is whether any of these types of concepts can live up to the constraints on a theory of concept-types per se. In § 2, I surveyed some motivations for Concept Atomism that amount to highlighting problems faced by Role/Structure theories; as a placeholder, I introduced the suggestion that role/structure features are hopelessly *messy*. We can now state this in terms of the claim that role/structure features are too messy to underwrite a theory of concept-types per se.

To sum up, the Concept Atomist can coherently embrace fine-grained types of concepts for the purposes of explaining various instances of FPP while denying that such types can play the theoretical role(s) demanded by a theory of concepts (concept-types). Indeed, while the atomist has reasons for denying that role/structure features play a direct role in the individuation of concept-types, the very characteristics that arguably frustrate Role/Structure theories of concept-types make role/structure features an attractive resource for the Concept Referentialist to bring to bear on FPP.

[40] When necessary, please treat regular quotes as corner quotes.

6 The persistence of FPP

This section draws attention to the heterogeneity and related instability of various instances of FPP and sums this up with the claim that FPP is very *persistent*. The persistence of the problem has implications for the polemical importance of FPP, including that it helps to vindicate two features of the approach endorsed in present paper that might otherwise raise suspicions. First, it suggests that it is a mistake to treat resolving FPP as a constraint on theories of concepts per se, which clears space for resolving FPP by appealing to non-concept-constitutive resources. Second, it helps to motivate the piecemeal strategy of bringing a catalogue of resources to bear on various cases, rather than holding out for a more unified solution.

Consider various potential extensions of the previous case involving Bill-the-Believer. Imagine that sometime after his trip to the grocery store, Bill comes across a packet of seeds labelled with both "Coriander" and "Cilantro" and pieces together that these are two names for the same plant. Bill's cognitive systems accommodate the new information; various tokens of the type CIL/COR IS CIL/COR that hitherto Bill would have refrained from believing, he is now willing to token as beliefs. Bill is no longer in the situation that was previously considered to be an instance of FPP. But now consider various ways in which Bill could end up in another situation that counts as an instance of FPP. Perhaps someone whom Bill trusts about such matters suggests to him that the seed packet in question contains seeds from two kinds of plants. Or perhaps Bill starts to worry that he has misconstrued an advertisement on the packet of seeds for an indication of its contents. These and other situations can lead to Bill's cognitive systems trafficking in tokens of CIL/COR that are treated as if they fail to co-refer and hence that lead to instances of FPP.

As well as there being various ways in which a case of FPP can arise and be resolved, existing cases can become more complicated. Go back to the original situation where Bill is in the grocery store. At this point Bill thinks (wrongly) that there are two types of plant at issue. Now imagine that Bill comes across a display of what initially he would describe as various "Cilantro" plants. Unbeknownst to Bill, the plants are from two different suppliers and despite being the same plant species, some look slightly different from the others and are labeled with the family name, "Apiaceae", rather than "Cilantro". As a result, Bill comes to think that what he had assumed to be one species of plant (remember, though, that he already assumes this is only one of two species) is in fact two different species. Bill now wrongly comes to believe that there are three species of plant at issue.

Along with the instability of specific cases, consider the heterogeneity of potential causes of FPP. Typical cases have their roots in the use of distinct linguistic expressions that an agent wrongly assumes do not co-refer. One point is that the expressions in question do not need to be distinct (e.g. "Hesperus" versus "Phosphorus"); the agent just needs to think that they are (e.g. multiple occurrences of "Jones" that an agent wrongly assumes to be names of different people). These beliefs can have various root causes: occurrences of the expression being pronounced differently, the expressions occurring in sufficiently different contexts (surely the Jones that works in the mailroom is not the Jones from the party), or something as esoteric as aberrant beliefs about how a single expression functions (e.g. that it refers in a context-sensi-

tive way) or how it is being used (e.g. as part of a one-time code). Moreover, arguably an agent doesn't need to *believe* that the expressions in question are distinct; they just need to worry that they *might* be.

Instances of FPP need not have their roots in linguistic expressions in the first place. Notice that what led to an instance of FPP in the previous discussion of Astra-the-Astronomer was the fact that Astra had various visual experiences that she failed to realize were experiences of the same object. Roughly speaking, an instance of FPP can be triggered by *anything* that has the potential to lead to various tokens of a concept having divergent role/structure features such that they are treated by cognitive systems as (potentially) referring to distinct individuals/kinds/properties.

One way to sum up the instability and heterogeneity of instances of FPP is in terms of what I will refer to as the *persistence* of the problem: given any specific catalogue of role/structure features, there will be room for further instances of FPP, namely those that turn on some additional difference in the role/structure features of the concepts in question that can lead to relevant cognitive systems (wrongly) treating various tokens of a concept as distinct.

One upshot of the persistence of FPP is that once one has committed to concepts being individuated in a way that is responsive to FPP,[41] it will be very difficult to specify well-defined identity conditions for such concepts; in particular, it is difficult to see how to avoid embracing concepts that are extremely fine-grained and arguably flexible.[42] In turn, it is difficult to see how to square flexible and fine-grained identity conditions with what are arguably among the most fundamental constraints on theories of concepts, namely that concepts need to be both compositional and sharable across agents, times, contexts, etc. This is, of course, a return to the issues sketched in § 2 regarding the difficulties presented by what I labelled the *messiness* of role/structure features of concepts. On the one hand, someone who treats FPP as a fundamental constraint on theories of concept individuation is likely to use this constraint as an argument against Concept Atomism and in favor of some or another Role/Structure theory of concepts. What I am now suggesting is that the Concept Atomist should use the persistence of FPP to turn this dialectic around, arguing that it will be difficult for any viable theory of concept individuation to fully resolve FPP without violating the even more fundamental constraints imposed by shareability and compositionality. Resolving FPP should not be accepted as a constraint on theories of concepts (concept-types) per se.

In contrast with the situation facing the Role-Structure theorist who tries to wield FPP against the Concept Atomist, the need for flexible and fine-grained resources fits very naturally with the approach to FPP articulated in the present paper. A key feature of the approach is to bring to bear resources that the Concept Atomist has reasons to treat as non-concept-constitutive. As such, the fact that role/structure features of concepts end up being hyper fine-grained and situation-specific is a virtue rather than

[41] A widely cited example of someone who appears to take on this commitment is Peacocke (1992).

[42] Rives (2009) makes a similar point, largely by citing Fodor (1998). Rives discusses so-called Mates' Cases (multiply-embedded attitude ascriptions) as lending independent support for the atomist's claim that whatever fully resolves Frege's Problem can't be semantic per se. My argument and conclusion differs in several respects, notably that I am applying the line of thought to concept individuation and that I think the issues are much more widespread and much more basic than relatively esoteric cases of iterated attitudes.

something that threatens the stability of the Concept Atomist's theory of concepts. Insofar as the Concept Atomist demotes role/structure features to non-concept-constitutive status, the persistence of FPP is easily accommodated; arguably it is what the Concept Atomist should predict based on the messiness of the features that can give rise to various instances of the problem.

The heterogeneity and instability of FPP also helps to vindicate what might otherwise seem like a theoretically messy, piecemeal approach advocated in this paper. The approach to FPP that I have advocated involves gesturing to a heterogenous catalogue of resources that can be brought to bear in potentially different ways on various putative problem cases. Given the persistence of FPP, this turns out to be a better fit for the shape of the problem than the assumption that there is a unified solution at a higher level of abstraction, for example one corresponding to such putative concepts as HESPERUS and PHOSPHORUS.[43]

7 Formal-types, mental files, and other 2-factor theories

I have resisted versions of atomism that embrace a concept-constitutive role for so-called *formal properties* and larger cognitive structures such as *mental files*—the kind of view that I earlier attributed to Fodor. This noted, following the discussion in § 5, I can now clarify that even the strong form of Concept Atomism to which I have cleaved is consistent with embracing not only these notions but a corresponding taxonomy (or taxonomies) of types of concepts. Once again, a crucial question is whether individuation in terms of such features can hope to live up to plausible constraints on a theory of concept-types. I think there are compelling reasons to be pessimistic about this. Since this is the main issue that drives me to reject less austere versions of Concept Atomism and to be more careful about how role/structure-features can play a role in resolving FPP without being concept-constitutive, it warrants further discussion.[44]

The role played by formal properties for a computationalist like Fodor—the reason for introducing formal properties in the first place—is that they are the properties to which cognitive processes are sensitive; in short, formal properties are introduced to ground facts about a concept's role in larger cognitive systems. Given that this is the motivation for positing relevant formal properties, it is not clear how individuation in terms of such properties can be in a better position than individuation in terms of role-features when it comes to underwriting types of concepts that are sufficiently robust (e.g. sharable and compositional). Fodor is in an especially awkward situa-

[43] I suspect the assumption that concepts need to be type-individuated at a level of abstraction corresponding to what is suggested by "HESPERUS" and "PHOSPHORUS" is an artifact of the kinds of cases that have occupied the limelight, and that this has made it natural to assume that a theory of concept-individuation can provide a general resolution to FPP.

[44] I focus on making a relatively basic point about the dialectical situation facing Fodor, one that pits core features of his overall approach against what I take to be relatively tentative attempts on his part to respond to Frege's Problem. For more in-depth developments of arguments in the vicinity of what I will sketch, see Aydede (1998), Onofri (2016), and Prinz (2002, Ch. 4). It is worth noting that Onofri explicitly intends the argument to extend to appeals to mental files.

tion here, given his repeated criticisms of accounts of concepts that appeal to role/structure features.[45]

Fodor (2008, ch. 3) also appeals to the notion of a *mental file*.[46] Unlike appeals to the formal properties of representations, I take mental files to be broader, architectural features of cognition. Still, it is not clear how appeals to files are in better shape than appeals to role/structure features with respect to the issues at hand. The way that the notion of mental files earns its keep comes down to the effects files are assumed to have on mental processes, for example that the contents of one file are (relatively) inferentially isolated from those of others. What I have summed up with a claim about the *persistence* of FPP suggests that if mental files are to underwrite a general resolution of FPP, file individuation would need to be determined by, or at least be able to track, potentially subtle, complicated, and easily shifted facts, for example regarding an agent's background beliefs, preferences, what information and which issues are salient in a particular context, and so on. Given the persistence of FPP and the messiness of potentially relevant role-features, I see no reason to expect that there will be a robust way to individuate anything that deserves to be identified as the Hesperus-file (as opposed to the Phosphorus-file), or the Superman-file (as opposed to the Clark-Kent file), and so on. As far as individuating concept-types is concerned, the situation with respect to file-structures is no better than it is for more direct appeals to role/structure features. In my view, mental files should be added to the list of features of the cognitive mind that the Concept Atomist can embrace and use to explain various features of cognition, potentially including various instances of FPP, but the Concept Atomist should make sure to do this in a way that does not require such features to play a direct role in the individuation of concept-types.

There are a host of other "2-factor" approaches that one can imagine being applied to concept individuation—if they haven't already. Obvious examples include Loar's (1985) distinction between *social content* and *psychological content* and Chalmers' (2011) *two dimensions* of content.[47] The basic pattern of my response to such views should be predictable: insofar as such second factors are tailor-made to resolve problems, such as FPP, that require flexible and fine-grained resources, they are likely to run into familiar constraints on a theory of concept-types. The Concept Atomist can be open-minded about the availability of these resources and their importance to

[45] There is a hint in Fodor's corpus of the realization that types grounded in formal properties are not sufficiently robust. When Fodor (1992a) invokes formal properties, he goes so far as to concede that he can envision giving up realism about types of psychological states (beliefs, desires, etc.) in spite of having an unshakable commitment to the reality of their intentional properties. The key issue comes down to whether the resources that Fodor wants to use to block substitution arguments in this context, namely formal properties, can underwrite sufficiently *robust* (my term) identity conditions for what in that context he refers to as types of psychological states. This discussion predates most of Fodor's published work on the nature of concepts per se, but I think the underlying moral generalizes, especially given Fodor's persistent claims about compositionality and shareability being key constraints on theories of concepts.

[46] Recanati (2010) provides a history of relevant uses of the notion of a mental file. According to Recanati, the key source of the idea is Perry (1980), who himself gives much credit to Donnellan's (1970, 1974) work on proper names.

[47] Past conversations with Loar suggest that he is in favor of a relatively liberal account of concept individuation that embraces both facts about referential content and facts about conceptual role.

explaining various aspects of cognition, but they should be wary about embracing the claim that they play a direct role in the individuation of concepts (concept-types).

8 The division of explanatory labor

A key move in this paper has been to embrace familiar resources but to demote them to non-concept-constitutive status. An obvious moral is that one needs to be careful when drawing conclusions about what the Concept Atomist can and cannot hope to explain based on the austerity of their theory of concepts per se. Theories of concepts are only part (although arguably somewhere near the core) of broader theoretical resources. As a result, it can be important to view an approach to concepts as going hand in hand with commitments (explicit or otherwise) regarding how best to *divide the explanatory labor* between a theory of concepts (concept-types) per se and other explanatory resources.[48]

The Concept Atomist has principled reasons for favoring the division of explanatory labor that goes along with their theory of concepts. In crude terms, concepts provide a relatively austere but at least robust (e.g. compositional and sharable) core for cognition; it then becomes natural for this core to be supplemented by various non-concept-constitutive resources. The complementary point is that the *messiness* of various non-concept-constitutive resources the atomist can hope to bring to bear on FPP is a good fit for what is demanded by the *persistence* of the problem. For the Concept Referentialist, the messiness of these resources is a good fit in another way: it is what one should expect given the role these features play in implementing the reference relation.

As outlined in § 4, I have argued that the Concept Referentialist can say something about the relationship between reference and potentially complicated nuts-and-bolts facts about the role/structure features of concepts. This helps to make sense of how instances of FPP can arise, as well as how role/structure features can be messy and yet play an explanatory role for the Concept Atomist in resolving FPP. Here it is worth noting, although this goes beyond anything covered in this paper, that this relationship can also help to make sense of how the (referential) content of concepts can play an explanatory role in psychological explanation despite cognitive processes being driven by distinct and more fine-grained features of concepts.[49]

[48] What I am describing as the *division of labor* is present in Margolis & Laurence (1999b), albeit not quite in the terms that I am using. The idea is captured succinctly by Brody & Feiman, who write the following: "Atomism simply denies that a concept's role in mental computation is constitutive of its identity. Far from being a disadvantage, this is one of the main benefits of embracing atomism: It provides an extremely clear division of labor between what concepts are and what concepts are used for" (Brody & Feiman, 2023, p. 15).

[49] The view of the relationship between reference and role/structure features that is articulated in § 4 is related to Fodor's (1994) picture of how an externalist theory of content and related view about intentional psychology can co-exist with a computational approach to mental processes. Fodor points out that this combination of commitments leaves room for two kinds of slippage: Frege Cases and Twin Cases. Having said this, I should acknowledge a worry, one raised by an anonymous referee, to the effect that insofar as the atomist cleaves to such an austere theory of concepts, they may have trouble making sense of the full range of explanatory generalizations that a theory of concepts might be expected to vindicate. Responding

Most of the discussion in this paper has been pitched in the relatively abstract terms of "role/structure features". I want to emphasize that the key points extend to more specific, empirically-supported claims about various kinds of cognitive structures and the processes in which they are implicated. I suspect that Fodor's often sharp treatment of psychological approaches to concepts has done much to fuel an assumption that the Concept Atomist is committed to dismissing the relevance of various kinds of cognitive structures that working cognitive psychologists and empirically oriented philosophers take to be closely related to the notion of a concept: prototypes, proxytypes (Prinz, 2002), exemplars, perceptual representations, constituents of theory-like structures, and so on. I want to end by endorsing a much more conciliatory picture. While the Concept Atomist has reasons for denying the specific claim that concepts can be straightforwardly identified with such structures, they can, and in my view should, take the available empirical evidence seriously and find a home for such structures in a more comprehensive account of the cognitive mind that has a relatively austere theory of concepts at its core.[50]

Acknowledgements Thanks to everyone over the past few decades who has tried to convince me either that the psychological analogue of Frege's Problem is the death knell for Concept Atomism or that the atomist has an easy solution. I especially need to thank Louise Anthony, Georges Rey, two very patient referees for this journal, and of course Jerry Fodor.

Declarations

Ethical approval N/A (my research does not involve human participants).

Conflict of interest No conflicts of interest to report.

References

Aydede, M. (1998). Fodor on concepts and Frege puzzles. *Pacific Philosophical Quarterly*, *79*(4), 289–294.
Bach, K. (2001). You don't say? *Synthese*, *128*(1/2), 15–44.
Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660.
Block, N. (1993). Holism, hyper-analyticity and hyper-compositionality. *Mind & Language*, *8*(1), 1–26.
Brody, G., & Feiman, R. (2023). Polysemy does not exist, at least not in the relevant sense. *Mind & Language*, mila.12474.

to this worry goes beyond what I can hope to accomplish in the present paper. For at least the start of how a referentialist can hope to respond to concerns about psychological explanation, see Schneider (2005).

[50] Kwong (2007) makes a similar but more narrow point about Concept Atomism and categorization.

Cappelen, H., & Lepore, E. (2005). *Insensitive semantics: A defense of semantic minimalism and speech act pluralism*. Blackwell.
Carey, S. (2009). *The origin of concepts*. Oxford University Press.
Chalmers, D. J. (2011). Propositions and attitude ascriptions: A Fregean account. *Nous*, *45*(4), 595–639.
Daley, K. (2010). The structure of lexical concepts. *Philosophical Studies*, *150*(3), 349–372.
Donnellan, K. (1970). Proper names and identifying descriptions. *Synthese*, *21*, 335–358.
Donnellan, K. (1974). Speaking of nothing. *Philosophical Review*, *83*(1), 3–31.
Edwards, K. (2009). What concepts do. *Synthese*, *170*(2), 289–310.
Edwards, K. (2010). Concept referentialism and the role of empty concepts. *Mind & Language*, *25*(1), 89–118.
Fodor, J. A. (1981a). Methodological solipsism considered as a research program in cognitive psychology. *RePresentations: Philosophical essays on the foundations of cognitive science*. MIT Press.
Fodor, J. A. (1992a). Substitution arguments and the individuation of belief. *A theory of content and other essays*. MIT Press.
Fodor, J. A. (1992b). *A theory of content and other essays*. MIT Press.
Fodor, J. A. (1994). *The elm and the expert: Mentalese and its semantics*. MIT Press.
Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford University Press.
Fodor, J. A. (2004). Having concepts: A brief refutation of the twentieth century. *Mind & Language*, *19*(1), 29–47.
Fodor, J. A. (2008). *LOT 2: The Language of thought revisited*. New York: Clarendon Press; Oxford University.
Fodor, J. A., & Lepore, E. (2002). *The compositionality papers*. Clarendon.
Fodor, J. A. (1981b). *Representations: Philosophical Essays on the Foundations of Cognitive Science* (1st MIT Press ed.). Cambridge, MA: MIT Press.
Fodor, J. A., Garrett, M. F., Walker, E. C. T., & Parkes, C. H. (1980). Against definitions. *Cognition*, *8*(3), 263–367.
Frege, G. (1948). [1892]: Sense and reference. *The Philosophical Review*, *57*(3), 209–230.
Hampton, J. A. (2006). Concepts as prototypes. *The Psychology of Learning and Motivation: Advances in Research and Theory*, *46*, 79–113.
Harman, G., & Greenberg, M. (2006). Conceptual role semantics. In E. Lepore, & B. Smith (Eds.), *The Oxford handbook of philosophy of Language*. Oxford University Press.
Kripke, S. (1979). A puzzle about belief. In A. Margalit (Ed.), *Meaning and use*. D. Reidel.
Kwong, J. M. C. (2007). Is conceptual atomism a plausible theory of concepts? *The Southern Journal of Philosophy*, *45*(3), 413–434.
Laurence, S., & Margolis, E. (1999a). Concepts and cognitive science. In E. Margolis, & S. Laurence (Eds.), *Concepts: Core readings*. MIT Press.
Laurence, S., & Margolis, E. (1999b). *Concepts: Core readings*. MIT Press.
Laurence, S., & Margolis, E. (2008). The ontology of concepts: Abstract objects or mental representations? *Noûs*, *41*(4), 561–593.
Laurence, S., & Margolis, E. (2020). Concepts. In E.N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2019 ed.).
Loar, B. (1985). Social content and psychological content. In R. H. Grimm, & D. D. Merrill (Eds.), *Contents of thought*. University of Arizona.
Machery, E. (2009). *Doing without concepts*. Oxford University Press.
Machery, E. (2010). Précis of doing without concepts. *Behavioral and Brain Sciences*, *33*(2-3), 195–206. discussion 206-144.
Millikan, R. G. (1989). Biosemantics. *The Journal of Philosophy*, *86*(6), 281–297.
Millikan, R. G. (2000). *On clear and confused ideas: An essay about substance concepts*. Cambridge University Press.
Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
Neander, K. (2017). *A mark of the mental: In defense of informational teleosemantics*. The MIT Press.
Onofri, A. (2016). Two constraints on a theory of concepts. *Dialectica*, *70*(1), 3–27.
Peacocke, C. (1992). *A study of concepts*. MIT Press.
Perry, J. (1980). A problem about continued belief. *Pacific Philosophical Quarterly*, *61*, 317–332.
Perry, J., & Blackburn, S. (1986). Thought without representation. *Proceedings of the Aristotelian Society, Supplementary Volumes, 60*, 137–166.
Piccinini, G., & Scott, S. (2006). Splitting concepts. *Philosophy of Science*, *73*(4), 390–409.
Prinz, J. (2002). *Furnishing the Mind: Concepts and their perceptual basis*. MIT Press.

Prinz, J. (2005). The return of concept empiricism. In H. Cohen, & C. Lefebvre (Eds.), *Handbook of categorization and cognitive science*. Elsevier.
Quilty-Dunn, J. (2021). Polysemy and thought: Toward a generative theory of concepts. *Mind & Language*, *36*(1), 158–185.
Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2023). The best game in town: The reemergence of the language-of-thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences*, *46*, e261.
Quine, W. V. O. (1970). *Philosophy of logic*. Prentice-Hall.
Recanati, F. (2010). Singular thought: In defence of acquaintance. In R. Jeshion (Ed.), *New essays on singular thought*. Oxford University Press.
Rives, B. (2009). Concept cartesianism, concept pragmatism, and Frege cases. *Philosophical Studies: an International Journal for Philosophy in the analytic Tradition*, *144*(2), 211–238.
Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology*, *104*, 192–233.
Russell, B. (2005) [1905]. On denoting. *Mind*, *114*(456), 873–887.
Schneider, S. (2005). Direct reference, psychological explanation, and Frege cases. *Mind & Language*, *20*(4), 423–447.
Schneider, S. (2011). *The language of thought: A new philosophical direction*. MIT Press.
Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Harvard University Press.
Vicente, A. (2010). Clusters: On the structure of lexical concepts. *Dialectica*, *64*(1), 79–106.
Vicente, A., & Martínez Manrique, F. (2016). The big concepts paper: A defence of hybridism. *The British Journal for the Philosophy of Science*, *67*(1), 59–88.
Weiskopf, D. A. (2009). The plurality of concepts. *Synthese*, *169*(1), 145–173.