Received: 14 May 2019 | Accepted: 18 June 2019

DOI: 10.1111/mila.12271

SPECIAL ISSUE ARTICLE

WILEY

Whither extensions?

David Pereplyotchik

Department of Philosophy, Kent State University (Kent Campus), Kent State University, Kent, Ohio

Correspondence
David Pereplyotchik, Department of Philosophy, Kent State University (Kent Campus), Kent State University, P.O. 5190, Kent, OH 44242-0001.
Email: dpereply@kent.edu

Abstract

Paul Pietroski develops an iconoclastic account of linguistic meaning. Here, I invite him to say more about what it implies about the relations between language, truth, and conceptual content. Readers concerned with securing the objectivity of conceptual thought may be worried about his claims that typical concepts "have no extensions" and that they "fit one another better than they fit the world." Others might applaud his anti-extensionalism in natural-language semantics but fear that his account re-raises familiar problems about extensions at the level of psychology.

KEYWORDS

classification, concepts, logic, meaning, reasoning, semantic internalism

1 | INTRODUCTION

Paul Pietroski's work constitutes genuine progress in semantics. *Conjoining Meanings* (*CM*) is a rich and iconoclastic book, jam-packed with convincing arguments against central pillars of received views, and many strikingly original positive proposals, motivated persuasively by historical, formal, and empirical considerations. Word limits are tight, so I am forced to be brief in praising this instructive, ambitious, and compelling contribution to the literature.

Pietroski's main goal in *CM* is to defend the hypothesis that linguistic meanings are "instructions for how to access and assemble concepts" (1). The "C-word" is important here, so Pietroski devotes a chapter to conceptual thought, and elaborates his views on conceptual content throughout. This is not, however, his central topic. Rather, the primary aim of *CM* is to develop and reinforce various positive proposals about *language*. Delving into the history of logic, troubles with liar-sentences, and event semantics, Pietroski patiently surveys a wide variety of semantic theories, classical and contemporary. Ultimately, he rejects Lewisian appeals to both "E-languages" and *possibilia*, as well as truth-theoretic semantics in its many guises, and develops a promising Chomskyan alternative.

No brief review can do justice to such an impressive array of topics. Commentators from formal semantics would focus on Pietroski's heretical counsels to abjure Fregean function application, Tarskian model theory, and other familiar technical machinery. Others might evaluate his proposals for reconstructing natural language semantics on a considerably slimmer footing. My own concerns will be more foundational. I will put Pietroski into dialogue with two kinds of "general" reader, each of whom approaches *CM* with a specific set of commitments and concerns about truth, language, and cognition. As already noted, Pietroski's leading semantic proposal (a) eschews appeal to truth—mostly—and (b) makes reference to concepts, construed as psychological entities with intentional contents. The readers I have in mind are particularly keen to learn what this conjunction implies about the relation between extensions and conceptual contents.

To facilitate the discussion, I will adopt a Platonic/Fodorian conceit, resurrecting two venerable opponents, each representing a familiar cluster of philosophical commitments. Rather than burdening any actual theorist (including myself) with either of their positions, I will let these fictional critics query Pietroski, narrating from the sidelines in the unindented text.

> *Hylas*: To my mind, extensional accounts of meaning and content are the cornerstones of a sound systematic philosophy. For broadly epistemological and metaphysical reasons, I am antecedently committed to *some* form of *Extensional/Truth-Conditional Possible-Worlds Semantics* (henceforth "PWS-ETC"), in either linguistics or psychology—ideally, both. Pietroski's arguments against PWS-ETC worry me. They threaten to undercut my views on incommensurability and my nonnegotiable commitment to Realism in science, ethics, and elsewhere. It was reassuring to me that his central hypothesis—namely, that "meanings are instructions for how to form truth-evaluable thoughts"—is compatible with the "insights" that motivated truth-theoretic semantics (68). But, while he says that *invented* languages "connect signals with extensions" (9), he argues that typical "natural" words and concepts have *no* extensions. His claim that concepts "fit one another better than they fit the world" (115) left me wondering what role extensions play in his view. Without a better idea of how Pietroski's semantic proposals anchor concepts to the world, I find his overall position difficult to evaluate.

> *Philonous*: To *my* mind, Pietroski's arguments against PWS-ETC are decisive, paving the way for semantic theories that can plausibly rival it in formal rigor, explanatory power, and philosophical motivation. I approached *CM* partial to one such theory—what I will label PEIRCianism: a Pragmatist Expressivism that takes Inferential Roles as Central to an account of conceptual/intentional content (cp. Brandom, 1994). But even PEIRCians need a subpersonal account of how concepts arise in real-time perception/comprehension, and Pietroski's is as good as I have seen. Nevertheless, although he rightly divorces extensions from linguistic *meaning*, some of his remarks seem to me to reintroduce them as explanatory posits at the level of conceptual/intentional *content*. In allowing that some concepts *have* extensions, and that some thoughts *can* be truth-evaluable, Pietroski leaves open metasemantic issues about what role extensions should play in cognitive science. But why think they have any use at all? I favor purging both linguistics *and* psychology of extensions. Theories that relegate them to psychosemantics strike us PEIRCians as simply "moving the bump in the rug". In reading *CM*, I often found myself wondering why Pietroski does not explicitly endorse a more thoroughgoing anti-extensionalism. If he eschewed talk of extensions in connection with

> both meaning *and* content, then his proposals about natural language semantics could be profitably integrated with the insights of PEIRCians—for example, deflationism about truth. This seems to me like a fruitful merger on several independent grounds.

Not all readers will have a stake in this debate. Psycholinguists interested in Pietroski's detailed accounts of constrained homophony data may see such big-picture issues as relatively peripheral to their concerns; ditto for semanticists evaluating, for example, his novel account of relative clause formation. But it is hard to imagine a *philosopher* with no commitments in this arena. On behalf of such readers, myself included, I will invite Pietroski to clarify and extend his position, and to say more about what implications he takes his overall account to have for the foundational issues that got so many philosophers interested in semantics in the first place.

2 | MEANINGS

The main negative contention of *CM* is that the notions of extension, truth, and denotation should play no explanatory role in a psychologically plausible semantics for natural languages ("Slangs"). Pietroski sees semantics as a naturalistic inquiry into "how Slang expressions are related to human concepts" (115). The best empirical theory of this psychological relation, he argues, does *not* identify meanings with extensions nor with determinants thereof. Although meanings facilitate the *assembly* of concepts, which *have* intentional contents, meanings are neither concepts nor their contents. Moreover, the relation between truth and conceptual/intentional content may likewise "turn out to be quite complicated and orthogonal to the central issues concerning how meanings compose."

Pietroski argues against PWS-ETC, viewing it as a widely accepted but nevertheless tendentious empirical hypothesis about Slang expressions.

> I do not think 'sky' is true of skies (or of sky), much less blue skies or night skies. I do not deny that there are chases, and that in this sense, chases exist even if skies do not. But the existence of chases does not show that 'chase' is true of them… [Likewise], there is no entity that 'Venice' denotes. In this respect, 'Venice' is like 'Vulcan', even though one can visit Venice but not Vulcan… I also agree that there is a sense in which there are blue skies, but no blue unicorns. But it does not follow that 'sky' is true of some things, at least not in the sense of 'true' that matters for a theory of truth… [T]here is no call to quantify over skies, in physics or linguistics. (68)

Importantly, Pietroski is *not* offering a metaphysical thesis about Venice, skies, or chases. In saying that "book" is like an empty expression in not having an extension, he is not intending to resuscitate anti-realism about, for example, books. Indeed, he says that "[t]he concepts BOOK:VEHICLE and BOOK:INFO apply to actual things, while UNICORN and GHOST do not." Nor is his claim merely that 'unicorn' and 'ghost' have *empty* extensions. Rather, "these words illustrate the general point that words don't *have* extensions."

In any event, extensions could not *be* meanings, if only because "expressions with different meanings can have the same 'extension'" (15). Fans of PWS-ETC appeal to "non-actual possibilities" in dealing with these issues, but Pietroski argues that this "is an odd way to maintain that meanings are extensions."

> If the meaning of a word is not whatever set of things that the word happens to be true of, why think the meaning is a mapping from each possible world *w* to whatever set of things that the word happens to be true of at *w*? [If] Slang expressions need not connect pronunciations to actual things, it seems contrived to insist that these expressions connect pronunciations to possible things… [I]nvoking possible unicorns is contrivance on stilts. (12)

Throughout *CM*, Pietroski develops other powerful arguments against PWS-ETC, detailing its failure to deal with liar-sentences, event descriptions, and other data—notably, the ubiquitous phenomenon of *polysemy*. Following Chomsky (2000), he points out that "water" is polysemously used to talk about various substances, many with lower H_2O contents than coffee, tea, and cola (21). Likewise, "France" can be used in expressing either of two "correct thoughts"—FRANCE:BORDER IS HEXAGONAL! and FRANCE:POLIS IS A REPUBLIC!.[1] But the polysemy of "France" "does not imply that something is both hexagonal and a republic, much less that 'France' denotes such a thing" (74). And, while "London" can be used to talk about "a particular location or a polis that could be relocated elsewhere," it is plain that "no location can be moved, and no political institution is a location." Pietroski concludes that "no entity is the denotation (or 'semantic value') of 'London'; the ordinary word has no denotation" (73).

On the alternative account that he goes on to develop, the lexical meaning of a polysemous expression "is an instruction for [fetching] a concept from a certain lexical address … shared by a *family* of concepts" (8). This raises a question about how hearers access the *relevant* concept(s) in a given context: What psychological mechanisms select just *one* of a "family" of concepts residing at a common lexical address? Philonous will eagerly note that this typically requires nondemonstrative pragmatic reasoning, of the kind he suspects is partly *constitutive* of conceptual/intentional contents. Hylas, for his part, will want to know more about whether concepts like FRANCE:BORDER and FRANCE:POLIS have extensions. If not, then are we forced into some sort of anti-realism about nation-states after all? To address such concerns, let us turn to our main topic—conceptual content.

3 | CONCEPTS

Concepts have two striking properties: "[1] they can be used to think about things in certain ways; [2] they exhibit significant relations of constituency, so that a [concept] used to think about things in one way (e.g., as cows) can be a part of another [concept] used to think about things in a distinct yet analytically related way (e.g., as brown cows)" (77). In connection with [2], Pietroski treats concepts as "mental symbols that are generated by languages of thought." Acknowledging that "some philosophers dislike talk of mental languages," he nevertheless adopts a "broad conception of languages as 'things' that connect interpretations of some kind with signifiers" (55) and sees "nothing wrong with the idea that certain biochemically realized procedures connect interpretations, often called *contents*, with head-internal signals" (85), whether in humans or bees. This raises foundational questions about what such "interpretations" are.

Pietroski's central thesis is that the interpretations of *Slang* expressions—that is, their *meanings*—are instructions ("*Begriffsplans*") for connecting pronunciations with concepts. Here, the relevant

[1] I respect Pietroski's notational conventions whenever possible, but I deliberately eschew hyphens when indicating Mentalese expressions that aren't *obviously* (mere) complex concepts (Section 4), and I also use "!" to indicate mental attitudes (e.g., realizing) and "⇒" for material inferential transitions.

connection between signifiers (Slang pronunciations) and interpretations (meanings/*Begriffsplans*) is mediated by a real-time psychological process—a "biochemically realized procedure"—to which pronunciations are the inputs. But if *Mentalese* signifiers (e.g., concepts or *Begriffsplans*) are "head-internal signals," then the notion of *their* interpretations—that is, their "*contents*"—is far less clear. Plainly, neither animals nor their brains need "biochemically realized procedures" for "connecting" *their own* concepts to "contents" or "interpretations"; it is not even clear what such interpretations would *be*.

> *Hylas*: It is clear to *me*. Presumably, *Begriffsplans* are the meanings/interpretations of pronunciations because they bear a suitable Naturalistic Relation (NR) to one another—perhaps, causal/nomic covariation or a historical/teleosemantic relation. Similarly, the interpretations of *Mentalese concepts* are whatever worldly denizens bear NR to *them*—that is, their *extensions*. So I can grant that PWS-ETC fails as a semantic theory of *Slang* expressions but still maintain that it succeeds when applied to certain *neural* episodes.
>
> *Philonous*: As a PEIRCian, I deny that there is a unique Naturalistic Relation between signifiers and their interpretations. Pietroski's idea that Slang meanings are instructions for concept assembly is independently attractive and requires no appeal to any such relation between pronunciations and meanings. More generally, I take "interpretations" to be the outputs of a cognitive interpretive process—for example, theorizing about how to paraphrase target signifiers (maps, bee dances, CPU states) into an already-understood metalanguage. In the case of Slang expressions, Pietroski is the theorist, pronunciations are the signifiers, and the relevant metalanguage is that of "composable instructions"—notably, *not* the language of truth conditions at possible worlds. When the signifiers are "head-internal signals," cognitive neuroscientists become the interpreters, using Slangs or specialized representational formats as their tentative metalanguage(s). In interpreting *neural* signifiers, they make free appeal to physiological, evolutionary, sociological, and psychofunctional considerations. For instance, if Pietroski is right that the structure of some complex concepts "mirrors" that of the corresponding *Begriffsplans* (25), then this constitutes a relevant kind of psychofunctional evidence. We PEIRCians hold that the downstream role of concepts in reasoning/planning and intentional action is another major source of evidence for hypotheses about how to best interpret neural/Mentalese signifiers.

Pietroski offers important clarifications about *his* notion of "content" in discussing how animals "think about the world." He emphasizes from the outset that, for the purposes of his project in *CM*, "[t]he relevant notion of thinking about is *intentional*." Concepts can be used "to think about things in [an] intentional sense that allows for wonder[ing] what unicorns like to eat… even though there are none… and for Hesperus-thoughts that are not Phosphorus-thoughts" (77). Thus, rather than investigating how concepts *succeed* in referring/denoting, Pietroski aims to accommodate potentially empty and coextenive-but-nonsynonymous concepts right from the get-go. Still, he makes a few key concessions to readers like Hylas.

> I grant that 'think about' can also be used to talk about a relation that thinkers can bear to entities/stuff to which concepts can apply; see, e.g., Burge. In this "*de re*" sense, one can think about bosons and dark

> matter only if the world includes bosons and dark matter, about which one can think. Any episode of thinking *de re* about Hesperus is an episode of thinking *de re* about Phosphorus. This extensional/externalistic notion has utility, in ordinary conversations and perhaps in cognitive science, when the task is to describe representations of a shared environment for an audience who may represent that environment differently. But however this notion is related to words like 'think' and 'about', many animals have concepts that let them think about things in ways that are individuated intentionally. We animals may also have representations that are more heavily anchored in reality; see, e.g., Pylyshyn. But a lot of thought is intentional, however we describe it. (80)

Hylas: I agree—we need to countenance intentional contents and representational purport more generally. But what about representational *success*? If neither meanings *nor* contents are extensions, then how can thinkers *know* about the world? I was relieved to learn that Pietroski's nonextensionalist semantics is compatible with the claim that "many concepts apply to mind-independent entities, like animals and planets, *where the notion of application is not intentional*" (78, italics mine). If there are *two* distinct notions of "application"—one intentional, and one not—then conceding the theoretical utility of the *non*intentional notion commits Pietroski to the existence of an "externalistic" mind-world relation. Accordingly, he acknowledges that "many conceptual contents are individuated externalistically; cp. Burge" (20) and that "[d]istinct concepts, which let thinkers think about things in different ways, may apply [nonintentionally] to the same things" (79). But, while this underwrites Pietroski's talk of "truth-evaluable thoughts" (68) and "misapplications" of concepts (20), he does not *specify* the relevant "externalistic" relation (NR). "Would encountering a few cows, or pictures of cows, suffice? Happily, my claims about meaning can be evaluated without settling these questions" (80). I disagree. Given that Pietroski's central thesis characterizes meanings in terms of concepts, which sometimes participate in *truth-evaluable thoughts*, I think he needs to say more about how concepts/thoughts latch onto the world.

Philonous: Hylas is demanding too much. (As ever.) By contrast, I applaud Pietroski for rejecting the common claim that so-called "*de re* thoughts" establish some sort of semantically relevant mind-world relation, perhaps on account of some special epistemic credentials. To his credit, Pietroski treats *de re*/*de dicto* constructions as (aspects of) thought *ascriptions*—that is, *not* as devices for picking out a privileged type of *thought*. Moreover, he properly circumscribes the function of such ascriptions to the navigation of perspectival divides (cp. Brandom, 1994). *De re* constructions allow speakers to describe the environment in which a creature is deploying concepts—as viewed *by the speaker* (and often her audience)—while *de dicto* constructions indicate *the concepts so deployed*.

Let us return to a related issue. Recall Hylas' earlier appeal to extensions in addressing incommensurability worries and explaining successful communication across perspectival divides. As Pietroski notes, "many theorists think that substantive disagreements about the things or stuff that we can talk about—for example, disagreements about the nature of rabbits, stars, or water—tell in favor of the idea that meanings determine extensions." But it is also clear that "people who know what 'star'

means can have different views about what things like Polaris and Sirius are: holes in the canopy; fires attached to the celestial sphere; effects of nuclear fusion…" (19).

> *Philonous*: To resolve this infamous tension, Pietroski endorses an inference that strikes me as particularly problematic: "[I]f speakers can use 'star' to advance very different hypotheses about the same things, this does suggest that speakers can connect 'star' with *distinct concepts that are extensionally equivalent*" (19, italics mine). This sounds like a substantial concession to Hylas. Elsewhere, though, Pietroski counsels us—wisely, in my view—to *give up* the idea that "successful communication requires speakers [to] use the same meanings/concepts." He sees it as a mere *idealization* that "members of a community have acquired the *same* language—or that they use the same words, or words that have the same meanings" (33).
>
> *Hylas*: This talk of idealization obscures Pietroski's acknowledgment that some concepts are extensionally equivalent, both within and across speakers. I was disappointed when he set aside questions about how different thinkers can apply *shared* concepts *correctly*: "Spelling out the details is hard, though not my job. My claim is that even if words like 'star' are initially paired with kind-concepts that have extensions, meanings don't determine extensions" (20). Although Pietroski *limits* "conceptual variation across (and within) speakers," by noting that "communities try to promote and maintain the external constraint," he ultimately concludes that it's merely "useful for some purposes" to "*pretend* that meanings are extensions, or that understanding is a matter of knowing what [Slang expressions] are true of". Without knowing more about which concepts have extensions, I am not sure what is involved in a community's "maintain[ing] the external constraint."
>
> *Philonous*: Pietroski is busy enough developing a path-breaking account of meanings and arguing—persuasively—that they are not extensions. What more do you want from the guy? Why think that he owes a story about some mythical mind-world relation?
>
> I agree with you, though, that *we* need an account of conceptual content, to supplement Pietroski's attractive account of how concepts are accessed/assembled by linguistic meanings. Access and assembly are, after all, just two of the things we do with concepts and both are far less exciting than, say, *reasoning* with them. As Pietroski emphasizes, "concepts that don't apply to anything can still play valuable cognitive roles (cp. Husserl, Dummett, Rey)" (79). Here, as elsewhere, he sounds an awful lot like a PEIRCian.

Philonous is onto something here, I think. But Pietroski's views on the nature of (subsentential) *classification/predication* and its relation to (sentence-sized) *thought content* appears to be incompatible with some core "PEIRCian" commitments. I want to now shift the discussion to those topics, which occupy the following two sections.

4 | CLASSIFICATION

The phrase "thinking about" is often used to ascribe cognitive *processes*, including global reasoning from current assumptions to new hypotheses/conclusions, plans, and actions (e.g., "We're thinking

about his tax returns"). But Pietroski follows the peculiar philosophical convention of using it to denote a punctate event of "classification."

> [C]oncepts have contents that can be described as ways of thinking about things; cf. Evans. A concept that can be used to think about something as a rabbit, whatever that amounts to, has a content that we can gesture at by talking about the concept type RABBIT. An instance of this type is a mental symbol that can be used to think about a rabbit as such, or to classify something—perhaps wrongly—as a rabbit; see Fodor. A concept of the type RABBIT-THAT-RAN, which can be used think about something as a rabbit that ran, is presumably a complex mental symbol whose constituents include an instance of RABBIT. A thought can be described as a sentential concept that lets us think about (some portion of) the universe as being a certain way. Thoughts of the type A-RABBIT-RAN can be used to think about the world as being such that a rabbit ran. (4)

Philonous: I assume that *thinking* is a matter of tokening *complete thoughts*—that is, intentional states that can be expressed (by linguistically competent creatures) only in *complete speech acts*, using *complete sentences*. Pietroski agrees that having a thought is tokening a "sentential concept." But he also says that *all* concepts are "ways of thinking about things," which puzzled me, given that *sub*sentential concepts are *not* complete thoughts. Tokening such a concept cannot *by itself* constitute "thinking about something"; to do that, subsentential concepts have to (in some way) "participate" in *sentential* ones. So while sentential concepts are "ways of thinking about things," *sub*sentential ones are better seen as *aspects of* such ways. Pietroski believes otherwise: "[H]earing 'Bessie' can… activate the denoter BESSIE, thereby leading [one] to think about Bessie in a singular way" (108). But it is not clear to me how activating the denoting concept BESSIE can *alone* constitute thinking about Bessie—in *any* way—even once. One can, of course, *stipulate* that thinking about things does not require tokening complete thoughts. But it is difficult to see what would motivate this. Perhaps introspection, but this is known to be an unreliable source of data, whether performed by naive speakers or by theoreticians.

These concerns transfer over to the notion of classification. I maintain that "classifying" is *not* the obtaining of some (super)natural relation between a concept and (a portion of) the actual world—let alone merely possible worlds. There are, instead, *acts* of classification—for example, asserting "That is a rabbit," or tokening the corresponding perceptual thought. Unfortunately, elliptical ascriptions of such acts can obscure their distinctively *propositional* contents in theoretically relevant ways.

In saying "Jamal classified some food as rabbit," a speaker (Juanita) purports to indicate some food—the *de re* component of the ascription—and then says what concept Jamal applied to it (viz., RABBIT). The word "as" marks the onset of the *de dicto* component, ensuring that Juanita does not commit *herself* to the correctness of Jamal's classification. (Perhaps, she knows that it is tofu.) Some theorists take it as a *datum* that RABBIT is the only concept Juanita ascribes to Jamal, and go on to conjecture that Jamal can deploy this subsentential concept *by itself* in classifying something. But that's implausible. If

> one can neither classify nor think about something without tokening complete thoughts, then Jamal's classificatory act requires tokening the complete thought, CCC IS-RABBIT!, where CCC is whatever concept he uses in thinking about the food.
>
> Pending psychological evidence to the contrary, we should say that subsentential concepts *play a role* in acts of classification—inner endorsements of *propositional* contents and the corresponding reports/assertions—but tokening a subsentential concept is *not* sufficient to carry off an act of classification.[2]

5 | SENTENCES

In articulating the structure of *Begriffsplans*, Pietroski posits six psycho-computational operations: two types of predicate conjunction ("M-junction" and "D-junction"), invention/introduction of concepts like AGENT(_, _), a limited operation of existential closure, concept polarization (⇑/⇓), and "M-abstraction"—a mental analogue of relative clause formation, like λ-abstraction, but "much less powerful… and limited to monadic concepts of a special sort" (109). Specifying this "special sort" raises questions about "the very idea of a matrix sentence."

Pietroski is skeptical that "Slangs generate sentences as such" and suspects that talk of "grammatical subjects" is just a roundabout way of "saying that tensed clauses have a 'left edge constituent' that somehow makes them complete sentences—whatever that amounts to—as opposed to mere phrases like 'telephoned Bingley'" (87). Rather than clarifying the notion of a "complete sentence," such talk *presupposes* it. Moreover, although "subject," "sentence," and related notions fund old-school subject-predicate conceptions of thought, they "may have no stable place" in contemporary scientific grammars (114).

> Linguists have replaced "S" with many phrase-like projections of functional items that include tense and agreement morphemes, along with various complementizers. This raises questions about what sentences are, and whether any grammatical notion corresponds to the notion of a truth-evaluable thought. But theories of grammatical structure—and to that extent, theories of the expressions that Slangs generate—have been improved by *not* positing a special category of sentence. So while such a category often plays a special role in the stipulations regarding invented formal languages, grammatical structure may be independent of any notion of sentence. (61)

Naturally, Pietroski does not characterize sentences—whether in Slang(s) or Mentalese(s)—by appealing to a distinction between sentential truth conditions and subsentential satisfaction conditions. Instead, he develops a novel version of Predicativism, according to which *all* concepts assembled by *Begriffsplans* are classificatory/predicative, including those assembled by sentence-sized Slang expressions. Pietroski divides predicative concepts into subsentential and sentential, positing mental operations that convert the former into the latter. "Given any concept M, applying the operation 'UP' yields a polarized concept, ⇑M, that applies to each thing if M applies to something" (30). For instance, if RABBIT applies to something, then ⇑RABBIT applies to each thing and ⇓RABBIT applies to no-thing.

[2] Nor is it obvious that "classifying" is a function of *all* concept application. Does *wondering* whether Bessie exists require classifying her?

To see how worries about "application" (Section 2) and "classification" (Section 3) intersect with Pietroski's treatment of sentential concepts, consider what distinguishes the sentential concept ⇑RABBIT from the *sub*sentential concept THING-IN-THE-WORLD.

> *Hylas*: Pietroski has the resources to distinguish the internal structures of these two concepts. As regards their content, he says that they can both correctly "apply to each thing in the world," and thus participate in true thoughts about the world. Since he countenances "extensionally equivalent" concepts and the explanatory utility of a *nonintentional* notion of application (Section 2), he should conclude that both ⇑RABBIT and THING-IN-THE-WORLD have *the world* as their extension, and leave it at that.
>
> *Philonous*: I disagree. The relevant notions of application/classification in this context are *intentional*. Pietroski was emphatic about this. And though he says that we can build truth-evaluable thoughts as a *side-effect* of language processing, he *denies* that "meanings are instructions for how to build concepts that exhibit classical semantic properties" (115). Likewise, he suspects that "most natural concepts [do not] have extensions; cp. Travis… if only because of vagueness; cp. Sainsbury" (9). Hence, *Begriffsplans* "make no reference to the things we usually think and talk about" (115). Moreover, because they have "mechanical execution conditions," it follows that "meanings [i.e., *Begriffsplans*] satisfy demanding compositionality constraints." Pietroski infers from this that the concepts thus assembled "'fit each other' better than they 'fit the world'"—a conclusion we PEIRCians find congenial. Still, his view that sentential concepts "apply to the world" raises large issues about how to characterize both The World—well lost?—and the *intentional* notions of application, classification, and predication. Let us revisit the latter.
>
> I assume that downstream from concept-assembly the endorsement of a "sentential concept" results in a complete *judgment*. It is not clear to me how Pietroski's ⇑/⇓ operators facilitate this—that is, how endorsement is superadded to polarized concepts after the *Begriffsplans* assemble them. If judgments and classifications are all "sentence-sized" (Section 3), then even *noticing* rabbits involves *judging that* they are in the relevant vicinity, or that The World "contains" them. Such judgments require deploying concepts *other* than RABBIT—for example, IS-HERE or IS-CONTAINED-IN-THE-WORLD. Pietroski says that in *applying* the concept THING-IN-THE-WORLD to things in the world, one *classifies* "the world" (or each "thing" in it). The same holds for ⇑RABBIT. But these classifications plainly differ in kind. No merely *formal* distinction can explain why THING-IN-THE-WORLD *cannot* fund a complete judgment, whereas ⇑RABBIT *can*. Moreover, if RABBIT is the only constituent of ⇑RABBIT, then it is not clear what judgment would result from endorsing ⇑RABBIT. A judgment about some *specific* $\text{rabbit}_{(\text{MASS-OR-COUNT})}$? (Which?) A classification of "the universe" (4) as "containing" it?

6 | LOGIC

Pietroski adds an important wrinkle in his discussion of distinctively *human* thinking, distinguishing between (a) *natural*, "relational" speech/thought, which takes a coarse-grained subject-predicate form, and (b) "categorical" thought, which tracks more fine-grained *logically valid* inferential relations. I quote at length, to convey the theoretical context in which this distinction arises.

> From a contemporary perspective, rules for agglomeratively adding and subtracting predicates seem trivial, at least as logic. But in thinking about the structure of human concepts, and our natural capacities to recognize instances of valid reasoning, the intuitive force of Aristotelian syllogisms is potentially relevant data. It seems that a wide range of classical inference patterns form a network that can be reduced to a few intuitively compelling principles, each of which may be getting close to conceptual bedrock… This suggests that predicates play a central role in a form of thought that is governed by a "natural logic" that treats predicate-reduction as a privileged form of inference… Frege assumed that we naturally think and talk in a subject-predicate format, and that we need help—[e.g.] his invented *Begriffsschrift*—in order to use our rudimentary capacities for relational thought in systematic ways… The idea was that a thought content can be "dimly grasped," in some natural way, and then re-presented in a more logically perspicuous format that highlights inferential relations to other contents… I think this is basically right: our categorical thoughts are governed by a natural logic that lets us appreciate certain implication relations among predicates; but our relational concepts are related in less systematic ways. We use relational concepts in natural modes of thought. But we do not reason with such concepts in the manner of an ideal Fregean thinker. (95–96)

> [N]atural sentences of type <t> may belong to languages of thought that are phylogenetically older than Slangs. Expressions of these newer languages may be used to build complex monadic concepts, perhaps including some special cases that are closely related to natural thoughts of type <t>. In which case, the very idea of a truth-conditional semantics for a human language may be fundamentally misguided. (114)

Hylas: Pietroski's view seems to be that, despite being "natural" *and truth-evaluable*, "relational" thoughts *do not* exhibit a structure that is relevant for recognizing fine-grained logical/analytic implication relations. By contrast, "categorical" thoughts are *not* truth-evaluable, but *do* mediate such recognition. This left me wondering how Pietroski views the relationship between natural-logical implication and *truth*.

Philonous: As a PEIRCian and a deflationist, I think recognizing the relevance of implication relations to intentional contents is a definite advance over "inflationary" theories that invoke truth as an *explanatory* notion in this context. But if we know that minds are *not* "ideal" (113) and that "[s]ubstantive principles of defeasible reasoning [are] more interesting" (95) than those of deductive logic, then why attend only to *logical* implication in plumbing "conceptual bedrock"? Plainly, *all* concepts—descriptive, modal, normative, etc.—play a role in inference/reasoning. Indeed, patterns of use are central to what concepts *are*—where "use" indicates ("non-solipsistic") functional roles, encompassing perceptual judgments, conceptually structured actions, valid logical reasoning, and, crucially, materially good global inferences. Brandom (2008) has argued persuasively that the classificatory and representational functions of concepts can be explained *in terms of* their inferential roles—though, notably, not vice versa. If he is right about this, then adopting some version of PEIRCianism would not only fit

smoothly with Pietroski's anti-extensionalist semantics, but would explain how even "empty" concepts can have contents, by capitalizing on their "valuable cognitive roles."

7 | REASONING

Pietroski recalls Kant and Frege's insistence that "a thought is not a mere list of ideas" (102). Notably, both philosophers were concerned with the role of concepts/thoughts in *reasoning*. Relatedly, both emphasized the fact that mental attitudes or forces—judgment, doubt, assertion, etc.—attach to *propositional* contents. So even if sentences are not distinguished by their having truth conditions, only things with *sentence*-sized contents can serve as premises/conclusions in reasoning, public or private.

If subsentential concepts cannot enter *directly* into inferences, then how do animals use them in, say, abduction or planning under uncertainty? In keeping with his language-of-thought picture, Pietroski endorses the *Computational Theory of Mind* (CTM), according to which cognition is driven *solely* by the structural features of quasi-linguistic representations that have their own syntax (Fodor, 1975). Tellingly, he points out that *Begriffsplans* can "take many forms… as in a von Neumann machine" (107).

Philonous: I agree that "exploiting analogies to computer[s]" renders the operative notion of instructions "tolerably clear and unobjectionable" (108). But this holds only for the *subpersonal* case, *not* for personal-level thought/inference (Pereplyotchik, 2017). After decades of GOFAI research, there is still no *general* classical computationalist account of nonmonotonic/abductive reasoning—the kind that requires global considerations of "simplicity" and "relevance." It would be helpful to learn how Pietroski navigates these issues, for he sometimes seems to suggest that *good* inferences must be formally licensed by operations that are sensitive solely to the formal structure of sentential concepts.

> [It] has long been clear that grammatical structure can diverge from the structure exhibited by the articulated thoughts that can be premises and conclusions of valid arguments… This raises the question of what makes good inferences good, and Aristotle's partial answer seems fundamentally correct. Premises and conclusions are sentences that exhibit a certain kind of structure, and particular arguments can be instances of valid forms of inference. (87)

CTM can be viewed as a species of "formalism," which seeks to explain the apparent goodness/validity of causal and material inferences—for example, THIS IS-LIGHTNING! ⇒ SOON, THERE WILL BE THUNDER! and THIS IS-A-DOG! ⇒ IT'S A-MAMMAL!—by treating them as enthymemes whose tacit premises are LIGHTING *ALWAYS*-FOLLOWS-THUNDER! and *ALL*-DOGS ARE-MAMMALS!, respectively. But CTM is known to face serious problems in accounting for what Fodor (1983) called "global" inferences. The argument, familiar from discussions of the frame problem (Fodor, 1987, 2002), has been transposed into a pragmatist key by Brandom (2008). The essential point is that the internal syntactic structure of mental representations is *local*, whereas nondemonstrative reasoning is sensitive to *global* properties/relations—for example, the relative simplicity of rival theories/hypotheses. In view of this, one must either drop CTM as an account of reasoning—pursuing, for example, connectionism (Churchland, 1990)—or find a way to circumscribe the information that a classical computational system must consider in completing a rational

task, as proponents of massive modularity suggest (Mercier and Sperber, 2017). The latter two proposals are in tension with LOTH, so Fodor rejected both. It would be valuable to hear how Pietroski handles these issues, given his tectonic commitments to both LOTH and CTM.

8 | CONCLUSION

In approaching a book that seeks to rid semantics of extensions, some reviewers would concentrate on the formal details of the dissenter's alternative proposal. I have largely neglected these, focusing instead on Pietroski's foundational assumptions throughout *CM*. In making ineliminable reference to concepts—that is, psychological entities that "have a certain content"—Pietroski's leading proposal about linguistic meaning reanimates much-discussed but never-settled metasemantic debates. I have argued that his commitments heavily constrain the shape that a metasemantics for concepts must ultimately take. Through Hylas and Philonous, I have invited him to elaborate his distinction between Mentalese signifiers that have extensions and those that do not. Hylas pressed Pietroski for an account of the "externalistic relation" that connects (some) concepts to their extensions. Philonous recommended PEIRCianism for both concepts and thoughts—those that *lack* extensions, and those that do not (if there be any). I will let our fictional critics round out the discussion.

Hylas: Pietroski rejects the view that "typical/natural" concepts "fit the world." And, following Chomsky, he maintains that "focusing on communication is not a useful way to get at the natural phenomenon of linguistic meaning" (33). But he also endorses (admittedly, in a footnote) a view of truth that seems to be at odds with these commitments.

> Davidson did not say that *sentences* are [true/false] relative to contexts. He held that certain spatiotemporally located *utterances*—roughly, those used to make assertions—are [true/false] (cp. Montague, Lewis, Kaplan). [F] or present purposes, we can ignore the distinction between instructions and utterances of instructions. My own view is that truth and falsity are properties of certain evaluative actions—for example, episodes of assertion or endorsement—and the corresponding propositional contents that are often described with the polysemous word "thought," as in "the thought that snow is white" or "the thought constructed by executing that instruction"; cp. Strawson, Soames. (63fn.)

This suggests that, *pace* deflationists, Pietroski takes truth be a real *explanatory* property of assertions/endorsements. I would be more comfortable with his "anti-extensionalist" semantics if I knew more about his views on *this* topic.

Philonous: Pietroski's central semantic thesis is *very* appealing, and his characterization of assertion/endorsement as "evaluative actions" is yet another point of convergence between him and PEIRCians. Still, I worry that his background assumptions about contents conflict with his general aversion to extensions. He emphasizes that even empty concepts can "play valuable cognitive roles," but he does not appeal to these in discussing their contents. Like Hylas, I want to hear more about what is it for linguistically accessed concepts to exhibit intentionality.

Some theorists attempt to furnish the mind and then—only as a *second* step—to determine how the furniture is typically *used* in communication and inference. I reject this two-step approach. Concepts are involved in memory/belief, learning, planning, and (closer to home) the interpretive processes that select one out of a family of co-lexical concepts in cases of polysemy. Pending an account of what such cognitive roles have to do with their intentional contents, Pietroski's view seems to suggest that *psycho*-semantics *must*, in the end, countenance extensions. Say it ain't so, Paul! Say it ain't so!

ACKNOWLEDGEMENTS

Thanks to Paul Pietroski, Daniel Harris, Ryan DeChant, and Will Fenton for helpful comments on earlier drafts.

ORCID

David Pereplyotchik https://orcid.org/0000-0002-4609-9574

REFERENCES

Brandom, R. (1994). *Making it explicit*. Cambridge, MA: Harvard University Press.

Brandom, R. (2008). *Between saying and doing*. Oxford: Oxford University Press.

Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge, MA: Cambridge University Press.

Churchland, P. (1990). *A neurocomputational perspective: The nature of mind and the structure of science*. Cambridge, MA: MIT Press.

Fodor, J. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Fodor, J. (1987). Modules, frames, and fridgeons. In J. Garfield (Ed.), *Modularity in knowledge representation and natural-language understanding* (pp. 25–36). Cambridge, MA: MIT Press.

Fodor, J. (2002). *The mind doesn't work that way*. Cambridge, MA: MIT Press.

Mercier, H. & Sperber, D. (2017). *The enigma of reason*. Cambridge, MA: Harvard University Press.

Pereplyotchik, D. (2017). *Psychosyntax: The nature of grammar and its place in the mind*. Cham: Springer International Publishers.

How to cite this article: Pereplyotchik D. Whither extensions? *Mind Lang*. 2019;1–14. https://doi.org/10.1111/mila.12271