Received: 10 January 2020 | Accepted: 24 March 2020

DOI: 10.1111/mila.12328

SOLICITED ARTICLE

WILEY

Polysemy and thought: Toward a generative theory of concepts

Jake Quilty-Dunn[1,2]

[1]Faculty of Philosophy, University of Oxford & Department of Philosophy, Washington University in St. Louis.

[2]Department of Philosophy, Washington University in St. Louis, One Brookings Drive, St. Louis, MO.

Correspondence
Jake Quilty-Dunn, Department of Philosophy, Washington University in St. Louis, One Brookings Drive, St. Louis, MO, 63130.
Email: quiltydunn@gmail.com

Funding information
H2020 European Research Council, Grant/Award Number: 681422

Most theories of concepts take concepts to be structured bodies of information used in categorization and inference. This paper argues for a version of atomism, on which concepts are unstructured symbols. However, traditional Fodorian atomism is falsified by polysemy and fails to provide an account of how concepts figure in cognition. This paper argues that concepts are generative pointers, that is, unstructured symbols that point to memory locations where cognitively useful bodies of information are stored and can be deployed to resolve polysemy. The notion of generative pointers allows for unresolved ambiguity in thought and provides a basis for conceptual engineering.

KEYWORDS
atomism, compositionality, concepts, essentialism, memory, polysemy

1 | INTRODUCTION

Concepts are mental representations of a special kind. While some mental representations are constitutively creatures of perception, language, action, or other domain-specific systems, concepts are the representational elements of thought. Many unsolved questions in the philosophy of mind and cognitive science concern the scope of concepts—such as: *How and when do children develop concepts? How do concepts influence perception? How do concepts mediate perception and action?* But a more fundamental unsolved question

concerns the nature of concepts themselves: *What is a concept?* Since concepts are elements of thoughts and thoughts are elements of reasoning, the answer to this question could hardly be more central to our understanding of human cognition. A theory of concepts is thus a core desideratum of current philosophy of mind and cognitive science.

Theories of concepts fall into two general varieties: those that construe concepts as unstructured symbols and those that construe concepts as structured bodies of information.[1] The first, atomistic sort of theory is typically motivated by a desire to explain what we might call the *compactness* of concepts: That is, the ease with which concepts are deployed in categorization and language comprehension and composed into complex structures. The second sort of theory is typically motivated by a desire to explain conceptual *richness*: That is, the highly articulated nature of the information that underwrites categorization and inference.

Both the emphasis on compactness and the atomistic theory that it motivates are canonically associated with Jerry Fodor (1998, 2004, 2008); see also Laurence & Margolis, 1999; Gleitman, Liberman, McLemore & Partee, 2019). A competing emphasis on richness has been far more prevalent in psychology and recent philosophy of cognitive science. Empirical research on the bodies of information that underlie categorization and inference has led theorists to posit variegated representational structures such as prototypes (Hampton, 1988; cf., Rosch, 1978),[2] exemplars (Medin & Schaffer, 1978), theories (Carey, 1985; Murphy, 2002), ideals (Barsalou, 1985), and essentialist beliefs (Gelman, 2003; Newman & Knobe, 2019). The heterogeneity of structures underlying categorization and inference raises the problem of whether concepts *qua* informational structures form natural kinds (Machery, 2009; cf., Weiskopf, 2009b; Vicente & Martínez Manrique, 2016).

This paper is an exploration of conceptual atomism. Atomism offers to provide an account of the compositionality of concepts while avoiding the messiness of heterogeneous informational structures. I will point out two problems with atomism, namely its failure to explain cognitive phenomena explained by informational structures and its failure to accommodate polysemy. I will then sketch a broadly atomistic theory that solves these problems. On the atomistic view defended below, concepts are *generative pointers*. Concepts are atomic symbols that (a) fail to determine denotations and (b) point to memory locations where informational structures are stored. The selection of stored informational structures modulates the content of episodes of thinking with a concept, thus allowing concepts to be polysemous and generate different denotations on different occasions.

[1]Some theorists deny that concepts are mental representations, instead taking them to be, for example, inferential abilities or Fregean senses (Evans, 1982; Peacocke, 1992). While there may be legitimate uses of the term "concept" captured by these theories, at least one legitimate usage refers to mental representations. Use of "concept" to refer to mental representations allows more direct contact with the scientific study of concepts and its historical roots in the "theory of ideas" predominant in both the rationalist and empiricist traditions in the 17th and 18th centuries (Fodor, 2003; Yolton, 1984). The aim of this paper is to articulate a theory of concepts constrained by this representationalist use of the term.

[2]It often goes unnoted that Rosch (1978) rejected a psychologistic notion of prototypes as mental representations, instead taking them to be (at most) abstractions over particular typicality judgments. Many later theorists, both friendly and critical of prototype theory, agree that the theory should be interpreted literally as a claim about representational structure (Fodor, 1998; Hampton, 1988; Smith & Medin, 1981).

2 | MOTIVATING ATOMISM

2.1 | Compositionality

The view that concepts are structured bodies of information is by far the dominant view among cognitive psychologists and philosophers of cognitive science (see Murphy, 2002, and Machery, 2009, respectively, for representative examples). The arguments for atomism are primarily negative: Theories that take concepts to be richly structured face unsurmountable problems. One basic problem concerns the compositionality of concepts. The hypothesis that concepts can compose together to form more complex representations offers to explain the productive and systematic nature of human thought (Chomsky, 1965; Fodor & Pylyshyn, 1988). However, the way concepts compose suggests that they cannot be identified with the structured bodies of information that underlie categorization and inference.

To illustrate this problem, I will focus on the view that concepts are prototypes (Hampton, 1988; Rosch, 1978; Smith & Medin, 1981). Prototypes are complex representations (cf., fn. 2) that specify features and how probable category instances are to have them. A prototype for CAT may include features like +FURRY, +FOUR LEGS, +TRIANGULAR EARS, and +MEOWS, all with quite high "weights." None of these features is necessary for being a cat, and one could imagine a creature that had all of them and yet was not a cat. Instead of being attuned to necessity and sufficiency, prototypes are attuned to typicality.[3] Since these features are typical of cats, and since any creature that possesses these features is probably a cat (i.e., the features have "high cue validity"), prototypes that include them are useful for categorization.

A major objection to prototype theory is the so-called "PET FISH problem" (Fodor, 1998; Fodor & Lepore, 2002b; Smith & Osherson, 1984). The PET FISH problem, in a nutshell, is that the concept PET FISH does not appear to be a function of the prototypes for PET (which might describe dog-like features) and FISH (which might describe trout-like features); indeed, the prototypical PET FISH is a goldfish, which is not particularly typical *qua* pet or *qua* fish. PET FISH is a complex concept that contains PET and FISH as constituents, and yet does not (so the objection goes) contain their corresponding prototypes as constituents. Therefore, the prototypes for PET and FISH are not constitutive of those concepts.

Prototype theorists have argued that prototypes can, in fact, compose. Models of prototype compositionality allow for modulation in light of background knowledge as well as acquisition and retrieval of prototypes based on experience rather than compositionality. Prinz (2012), for example, proposes a three-stage "RCA" model consisting of retrieval, composition, and analysis. *Retrieval*, the first stage, involves searching for stored prototypes: "When we are given two concepts to combine, we first search memory for relevant knowledge. In some cases, we will have stored concepts corresponding to the compound" (Prinz, 2012, p. 448). In a case such as this (e.g., PET FISH), we happen to have a stored prototype "cross-listed" that can be retrieved (e.g., a prototype that matches goldfish).[4] If there is no such information to be retrieved, then we engage in

[3]There are prototype-like structures that are not attuned to typicality, such as "ideals," which are attuned to normativity—for example, while the typical pub may not be particularly cozy, the *ideal* pub is (Barsalou, 1985). Del Pinal (2016) argues that prototypes may be attuned to more than mere typicality and can encode abstract features as well as dependency relations between features. However, Del Pinal arguably lumps distinct representational structures (such as ideals and causal models/theories) under the single label "prototype."

[4]Prinz's model in fact appeals to stored exemplars and on-the-fly construction of prototypes (see also Barsalou, 1987). I assume here that prototypes themselves can be stored in long-term memory (Hampton, 2015; Murphy, 2016).

composition by pooling features and weights together from the constituent prototypes (Hampton, 1991). Finally, through *analysis* we modulate the weights and features in the newly constructed prototype in light of "background information" (Prinz, 2012, p. 448). Prinz's RCA model is really a sketch that captures the essence of more specific proposals for prototype compositionality (Hampton, 1991; Hampton & Jönsson, 2012), including in positing an initial retrieval stage and a final stage of analysis (Prinz, 2012, p. 449).

I grant for the sake of argument that RCA models are empirically adequate. Our present interest is instead whether they preserve *the hypothesis that concepts are prototypes*. They do not appear to do so. Consider what RCA models posit as the initial, mandatory stage of thinking with complex concepts: Searching long-term memory for information proprietary to the complex concept. The logic of this stage coming first entails that, when we grasp the meaning of "pet fish," we are thereby able to search for PET FISH information in long-term memory without first combining prototypes. But this requires that the ability to deploy the complex concept PET FISH is prior to the composition of prototypes.[5] Indeed, the model predicts that we only go on to compose prototypes "[i]f the retrieval stage bears no fruit" (Prinz, 2012, p. 448). Tokening PET FISH is therefore independent of prototype compositionality. One could try to switch the order around and posit a "CAR" model on which composing prototypes occurs prior to retrieval. However, this model would require, implausibly, that we compose the prototypes for PET and FISH every time we think PET FISH and then throw the resulting structure away once we retrieve the stored goldfish-tracking prototype (cf., Del Pinal, 2016).[6]

In short, RCA models provide a plausible account of prototype compositionality at the expense of the prototype theory of concepts. Prinz perhaps accepts this consequence, appealing to the heterogeneity of informational structures: "[T]he question about prototypes is not *whether* they are concepts but *when* they are concepts" (Prinz, 2012, p. 440; see also Hampton, 2010). The PET FISH problem suggests that complex concepts should never be identified with prototypes. Moreover, the heterogeneity of informational structures only exacerbates the problem. Theories, exemplars, and other structures are no better equipped to compose than prototypes.[7] The hypothesis that concepts are not informational structures at all, but are instead unstructured atomic symbols, avoids the PET FISH problem. Composing FISH into PET FISH does not bring along the features that FISH brings to mind; those features are not parts of FISH, which lacks any internal structure whatsoever.

[5]One could deny that understanding "pet fish" requires deploying PET or FISH. But this reply simply denies that phrasal meanings have internal semantic structure and thus rejects compositionality outright, assimilating all phrasal meanings to noncompositional idioms like "kick the bucket" (Canal, Pesciarelli, Vespignani, Molinaro & Cacciari, 2017; Peterson, Burgess, Dell & Eberhard, 2001). Even many *idioms* have internal semantic structure, however, such as "spill the beans," where "beans" denotes information and "spill" denotes divulging (Nunberg, Sag & Wasow, 1994; Titone & Connine, 1999). It should serve as a theory-neutral datum that grasping nonidiomatic phrasal meanings requires composing lexical meanings.

[6]Prinz (2002, pp. 291–295) and Weiskopf (2009a) argue that the PET FISH problem conflates the plausible constraint that concepts must *be able to* compose with the implausible constraint that they *always do* compose. But the problem here is that the composition of PET and FISH required to grasp PET FISH is not prototype composition.

[7]For example, PET FISH exemplars are not derived from PET and FISH exemplars, and while a theory for WATER might specify that *water* = H_2O, that theory does not compose into SWAMP WATER (Malt, 1994). These theories also face independent problems (Prinz, 2002, pp. 75–88; Fodor, 1998, pp. 112–119; Murphy, 2016).

2.2 | Messiness

The heterogeneity of informational structures creates an additional problem for the thesis that concepts are structured bodies of information: Which informational structures constitute a particular concept? Prototypes, exemplars, and theoretical models under a particular concept are not all deployed together (Machery, 2009). Informational structures are *messy*; it is unclear how to draw a line around some pattern of activation and call it the concept FISH.

Prinz (2011) gives voice to an increasingly popular ecumenicist attitude: "[M]any concepts comprise the full range of representations used to categorize, including prototypes, exemplars, folk theories, and so on" (p. 2). Pursuing one elaboration of Prinzian ecumenicism, Vicente and Martínez Manrique (2016; "VMM") defend a "big concepts" view, on which the whole bundle of structures constitutes a single hybrid concept. Hybridism incurs a burden of articulating the psychological principle that unifies these structures. For VMM, that principle is *functional stable coactivation*: Structures fall under the same concept when they activate each other in a way that facilitates task performance stably across contexts and tasks.

This principle is problematic, however, given the ubiquity of spreading activation across distinct concepts. Since CAT and DOG are associated, they activate each other. While there may be some instability in semantic priming across contexts (Stolz, Besner & Carr, 2005), there are contextual effects on priming under a single concept as well (Barsalou, 1982). Activation within a cluster of informational structures need not be qualitatively more stable than strong associative links like CAT–DOG. VMM also appeal to speed of activation. But across-concept activation is likely at least as quick as activation within "big concepts." Masked pictures of animals presented for 13 ms successfully prime animal-related words, suggesting that activation spreads from CAT to DOG nearly instantaneously (Van den Bussche, Notebaert & Reynvoet, 2009).

Moreover, associative links across concepts can be stably functional. VMM approvingly cite Anderson's (1983) ACT model of spreading activation. On Anderson's model, however, an activated node in a network activates all nodes directly connected to it. In that case, activation that spreads from DOG to CAT will "reverberate" backward from CAT to DOG (Anderson, 1980, p. 265). As McNamara puts it, "[a]ctivation spreads from the prime to the target, from the target to the prime, and back again, until a stable pattern of activation is reached" (McNamara, 1992, p. 1177). Activating CAT in response to DOG will thus be functional, since the reverberatory activation back to DOG will facilitate the use of DOG. And since activation facilitates use in a domain-general way (e.g., activation facilitates not only linguistic processing but also visual categorization—Sperber et al., 1979), the coactivation of CAT with DOG is not only functional but also stable across tasks. Functional stable coactivation thus cannot be the principle that unifies informational structures. Hybridism fails to avoid the problem of messiness.

Weiskopf (2009b) defends another ecumenist view, on which each informational structure constitutes a particular instance of a concept. According to this pluralist thesis, CAT constitutes a superordinate kind of which the prototype, theory, and so forth are subkinds. Pluralism faces the same problem of messiness. Weiskopf appeals to "identity links" (Weiskopf, 2009b, p. 166) that connect informational structures. For Weiskopf, identity links might be realized in "semantic networks" (Weiskopf, 2009b), but as we have seen, activation links in semantic networks fail to distinguish information that falls under a concept from merely associated concepts. The presence of an explicit identity statement also seems insufficient, since we might believe that A = B but have distinct concepts for A and B. Lois Lane might know that Clark Kent = Superman but still retain distinct concepts, as demonstrated by her ability to think Frege-like thoughts without

contradiction, such as JIMMY OLSON KNOWS HE WORKS WITH CLARK KENT BUT DOESN'T KNOW HE WORKS WITH SUPERMAN.

Atomism avoids the problem of messiness by positing a single unstructured symbol that is tokened whenever the concept is tokened, functioning like a tag that psychologically marks the concept. Moreover, on atomism, cognitive processes such as composition and logical inference can be run over concepts independently of background information (Weiskopf, 2009b). Atomism offers to provide a step toward solving the problem of messiness: The information that falls under a concept is the information that is appropriately functionally related to this unstructured symbol. How this functional relation should be understood, however, is just one of the outstanding problems for atomism.

3 | PROBLEMATIZING ATOMISM

3.1 | Atomism, inference, and lexical meaning

Atomism can be understood as making a claim about both the syntax and the semantics of lexical concepts: Syntactically, they lack internal structure, and semantically, they fix denotations. Thus concepts for the atomist are "amodal, unstructured symbols that represent determinate referents" (Prinz, 2011, p. 16).

The syntactic thesis raises a problem about how atomistic concepts help explain human cognition. When we use DOG to think about dogs, we do not just idly fixate on doghood; we may draw inferences about things like animalhood, having four legs, barking, begetting dogs, and other properties represented in our informational structures. One problem for atomism is to explain how concepts provide rich information for cognition while lacking any internal structure.

The semantic thesis, on the other hand, runs into an entirely independent problem: Concepts do not represent determinate referents. The primary evidence for this claim comes from the phenomenon of *polysemy* (Apresjan, 1974; Falkum, 2015; Machery & Seppälä, 2011; Nunberg, 1979; Pustejovsky, 1995; Vicente, 2018). Polysemy is a form of lexical ambiguity. Lexical ambiguity occurs when a single ortho-phonological wordform can be used to express distinct denotations. For example, "bank" can refer to a financial bank or a riverbank. This form of ambiguity, in which distinct meanings seem to be completely independent, is *homonymy*. Distinct meanings of a homonymous word cannot be used together:

(1) #The bank cashes checks and slopes into the river.

There is no reading of (1) on which the occurrence of "bank" in (1) refers both to a financial bank and a riverbank. Any coherent reading would require some outlandish backstory (e.g., an animate, resourceful riverbank, or a creative architect). Homonyms are distinct words that happen to share orthography/phonology.

Not all lexical ambiguity is homonymy. The word "bottle," for example, is ambiguous between a container ("Mary held a bottle of beer") and what it contains ("Mary drank two bottles of beer"). But these distinct denotations can be related via anaphoric binding:

(2) Mary quickly drank her bottle of beer and then smashed it on the floor.

The thing Mary drank is liquid, and presumably in her stomach, and the thing she smashed is glass, and presumably on the floor; yet the anaphor "it" is successfully dependent on "bottle" despite the shift in denotation. This sort of flexibility is characteristic of polysemy and is impermissible for homonyms. There are counterexamples—for instance, the meat-versus-animal senses of "lamb" do not admit of anaphoric binding or the related phenomenon of copredication (Ortega Andrés & Vicente, 2019). As I will argue below, the best evidence for the polysemy/homonymy distinction is experimental. Nonetheless, contrasts like that between (1) and (2) provide an intuitive foothold into the distinction.

There is no consensus on how to understand polysemy. But a common hypothesis is that polysemous words have a single meaning that can be modulated to fix distinct denotations depending on context (Pustejovsky, 1995).[8] While homonyms have distinct lexical entries, polysemes are mapped to a single lexical entry with a single meaning that enables access to multiple distinct senses. Given the standard assumption (accepted by Fodor, among many others) that word meanings are represented by concepts, it follows that polysemy undermines the atomistic idea that each concept has a single denotation. Instead, concepts and word meanings shift their denotations.

If concepts are structured bodies of information, then shifts in denotation can arise from deploying particular bits of information rather than others. Vicente and Martínez Manrique make this move on behalf of hybridism (2016, pp. 81ff). The lack of internal structure in atomistic concepts means atomism struggles to account for polysemy. Hampton sums up this point concisely:

> Words may be radically polysemous (Nunberg, 1979). If concepts are to be tied fairly closely to substantive words (as just about everyone, including Fodor, would have them be) then concepts too must be amenable to many and varied contributions as components of thoughts. How this is possible without some kind of internal structure is problematic. (Hampton, 2000, p. 302)

In response, Fodor (1998) denies that the polysemy/homonymy distinction tracks anything of semantic significance—or, put more defiantly, that "there is no such thing as polysemy" (p. 53). His strategy is to slot putative cases of polysemy into one of two categories: either (a) they fail to be genuinely ambiguous, or (b) they are really just homonymous.

An example of the first category is the word "keep." While Jackendoff (1992) argues that "keep" has distinct senses in "keep your money" and "keep your job," Fodor argues that "keep" is univocal:

> People sometimes used to say that "exist" must be ambiguous because look at the difference between "chairs exist" and "numbers exist." A familiar reply goes: the difference between the existence of chairs and the existence of numbers seems, on reflection, strikingly like the difference between numbers and chairs. (Fodor, 1998, p. 54)

[8]The nature of "context" is opaque. One might take it to be fixed by the content of a discourse representation (Heim, 1982), as common ground (Stalnaker, 2014), or as some set of extra-mental facts relevant to the fixation of meaning (Kaplan, 1978). I use the term liberally here, to include virtually all facts relevant to the modulation of a word meaning, including pragmatic factors (Carston, This volume).

In other cases, however, Fodor grants lexical ambiguity. For example, Fodor and Lepore grant that "bake" has distinct senses in "bake a cake" and "bake a potato." The former implies an act of creation (e.g., in a mixing bowl) while the latter simply involves heating up. Fodor and Lepore (2002a) accept that "[a]pparent polysemy is generally real; the reason 'bake' seems to be lexically ambiguous is that it is" (p. 110). More straightforward cases of genuinely ambiguous polysemes include moves from an animal to its meat (e.g., "lamb"). Fodor and Lepore assimilate such cases to homonymy:

> Surely there just *couldn't be* a word that's polysemous between *lamb-the-animal* and (say) *beef-the-meat*? Or between *lamb-the-animal* and *succotash-the-mixed-vegetable*? That there couldn't may itself sound like a deep fact of lexical semantics. But no; it's just the truism that, the less one can see what the relation between X and Y might be, the more one is likely to think of an expression that is X/Y ambiguous as homonymous rather than polysemous. (Fodor & Lepore, 2002a, p. 117)

For Fodor and Lepore, the lexicon does not distinguish polysemy from homonymy. Ambiguous expressions are always mapped to multiple distinct lexical entries. The distinction consists instead in the fact that language users think of polysemes as related. Thus the distinction between polysemy and homonymy is *metalinguistic*, not semantic. In that case, polysemy provides no evidence against the atomistic view that concepts have unique referents.

I will now argue that this strategy fails. Polysemous expressions are distinguished *within the lexicon* from mere homonyms; they involve a single lexical entry with a single concept. I will then consider a defensive move on behalf of atomism—namely, that concepts are not the representations used to grasp word meanings, and hence that polysemy is compatible with a denotational conceptual atomism—and argue that it is unmotivated.

3.2 | Polysemy in the lexicon

The debate between denotational semanticists like Fodor and Lepore and nondenotational semanticists like Pustejovsky often rested on intuitions. Is the same meaning of "lamb" involved in "X hugged a lamb" and "X ate some lamb?" Some intuitions may be more robust, such as copredication and anaphoric-binding tests for polysemy (Ortega Andrés & Vicente, 2019). But such tests fail for some polysemes (e.g., "lamb"). It is hard to show through intuitions alone that a polysemous expression has a single lexical entry.

Fortunately, in the two decades since Pustejovsky's exchange with Fodor and Lepore, there has been substantial experimental work on how lexical items are processed, including polysemous and homonymous ones. This work strongly points toward the hypothesis that homonymous expressions involve multiple distinct lexical entries and polysemous ones do not—call this *the single-entry hypothesis*. I focus on two strands of evidence: priming/reaction-time evidence and developmental evidence.

3.2.1 | Priming and reaction-time evidence

The single-entry hypothesis holds that resolving lexical ambiguity works in two different ways. In homonymy, one meaning is *selected* among multiple lexical entries. In polysemy, a common

semantic representation is *modulated* within a single lexical entry. This difference between selection and modulation furnishes some concrete empirical predictions. Given the competitive nature of selection, we would expect homonymous meanings to inhibit rather than prime each other. Meanwhile, since polysemous senses are accessed via a single lexical entry, those senses should strengthen and prime each other. We should also expect the larger amount of stored information to facilitate retrieval of polysemes, since reverberatory priming amongst this information should allow the polyseme to reach a critical threshold of activation more easily. Thus we should expect to find enhanced priming for polysemous words, and disruption for homonymous words: minimally, failures of priming, but also inhibition and increased duration of meaning retrieval.

A well-known effect in the psycholinguistics literature is the "ambiguity advantage": Ambiguous words are processed more quickly than unambiguous words (Kawamoto, Farrar & Kello, 1994; Rubenstein, Garfield & Millikan, 1970). However, psycholinguists in previous decades regularly conflated homonymy from polysemy, even confusingly using "polysemy" to refer to both cases (Marcel, 1980). Klepousniotou and Baum (2007) carefully separated homonymous words from polysemous words and found that the advantage occurred only for polysemous words. Indeed, homonymous words trended toward being worse than controls (also found by Rodd, Gaskell & Marslen-Wilson, 2002); Maciejewski and Klepousniotou (2020) later found a robust "ambiguity disadvantage" for balanced homonyms (i.e., homonyms of roughly equal frequency). The ambiguity advantage for polysemous words is predicted by the single-entry hypothesis.

Frazier and Rayner (1990) used eye tracking to tell how long subjects looked at ambiguous words. They distinguished polysemous from homonymous words and embedded the words in a sentential context that disambiguated their meanings; this disambiguating context came either before or after the ambiguous word. For example, "dinner" can denote a meal or an event. Sentential disambiguation can occur before ("Tasting burned[/Ending early], the dinner wasn't very enjoyable") or after ("Apparently the dinner wasn't very enjoyable, tasting burned[/ending early]"). When disambiguating information came after, reading time was slowed for homonymous words (e.g., "It seems that the suit bothered Dick, wrinkling so easily[/progressing so slowly]"), but *not for polysemous words*. This kind of asymmetry is precisely what the single-entry hypothesis predicts. Reading a homonymous word requires selecting among multiple lexical entries and reading a polysemous word does not; thus a lack of disambiguating context disrupts reading for homonymy but not for polysemy (see also Frisson & Pickering, 1998; Brocher, Foraker & Koenig, 2016; see Eddington & Tokowicz, 2015 for a review).[9]

One of the main pieces of psycholinguistic evidence against the single-entry hypothesis comes from Klein and Murphy (2001). They found no polysemy/homonymy distinction in a priming experiment, using polysemous words like "paper" (material vs. newspaper/institution). They used adjectives to disambiguate a particular sense ("shredded paper") and found priming for adjectival phrases with the same sense ("wrapping paper") and, crucially, inhibition for phrases with an inconsistent sense ("liberal paper"), which is also found for homonyms. This result is at odds with the single-entry hypothesis. However, Klepousniotou et al. (2008) took

[9]Shen and Li (2016) reported homonymy-like disambiguation effects on reading time for polysemous words in Chinese. However, Klepousniotou, Titone and Romero (2008) found characteristic priming effects only for high-overlap words, which had a mean similarity rating of 4.12 out of 5; in contrast, the mean for Shen and Li's "polysemous" stimuli was 4.19 out of 7, more akin to Klepousniotou et al.'s moderate overlap stimuli (2.97 out of 5). Thus their results (like Klein & Murphy, 2001) arguably reflect inadequately polysemous materials.

care to distinguish high-overlap from moderate-overlap (e.g., metaphorical) and low-overlap (homonymous) senses and found a polysemy/homonymy distinction on Klein and Murphy's task for high-overlap stimuli.

Neurolinguistics provides relevant evidence as well. Klepousniotou, Pike, Steinhauer and Gracco (2012) used electroencephalography to measure the N400, which is a negative-trending event-related potential that signals semantic expectation violation. For example, if you read a sentence like "The boy played fetch outside with his," the expectation is that the word "dog" will appear. If instead "lizard" appears, then after ∼400 ms, the N400 signal will be increased (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980). Klepousniotou et al. examined the N400 across dominant and subordinate senses of polysemous words. For example, while "rabbit" is polysemous between an animal and its meat, the former is typically dominant. Similarly, "bank" is homonymous but the financial meaning is typically dominant (*modulo* context—Rodd et al., 2016). For homonyms, the N400 is reduced for words related to dominant meanings (Maciejewski & Klepousniotou, 2020); for polysemous words, however, the N400 is reduced for both dominant and subordinate senses. Even when context activates a dominant sense of a polyseme, the subordinate sense remains primed. This sort of evidence is exactly what the single-entry hypothesis predicts: Resolution of polysemous words primes rather than inhibits other senses because they are accessed via the same lexical entry.

Later, MacGregor, Bouwsema and Klepousniotou (2015) showed that the N400 could be found even at long intervals (750 ms) between polysemous primes and targets, while the N400 was only found for homonyms at shorter intervals. This result suggests that polysemous senses prime and strengthen each other, allowing lasting activation patterns, while homonymous meanings compete for selection (Maciejewski & Klepousniotou, 2020), causing inhibition and quicker decay of activation (Vicente, 2018).

In short, psycholinguistic and neurolinguistic evidence suggests that polysemous words, unlike homonyms, are easier and quicker to retrieve and that ambiguity resolution for polysemes involves modulation of a common meaning rather than selection among competing concepts. Thus polysemy, unlike homynymy, involves a single concept with multiple available denotations.

3.2.2 | Polysemy in development

An independent source of evidence for the single-entry hypothesis comes from developmental psychology. Recall that, for Fodor and Lepore, the distinction between homonymy and polysemy is fundamentally metalinguistic—all forms of ambiguity involve a proliferation of lexical entries, and polysemy consists merely in a metalinguistic sense that some ambiguous words are related. Children therefore make an interesting test case given their general lack of metalinguistic competence. Children under (roughly) age seven regularly fail metalinguistic tests. For example, if asked to list the number of words in a sentence, young children instead list the number of objects or events (Bialystock, 1986). In multiple experiments, Srinivasan and Snedeker (2011, 2014) took advantage of the metalinguistic incompetence of 4-year-olds by seeing whether they nonetheless show sensitivity to the polysemy/homonymy distinction. If they do, it is unlikely to be metalinguistic; instead, it likely arises out of the lexicon itself, as the single-entry hypothesis contends.

Srinivasan and Snedeker (2011) taught children a novel word in a context that mapped it to one sense of a polysemous word. Elmo from *Sesame Street* talked about his "blicket," which is

red and fits in his backpack, thus mapping "blicket" to the physical-token sense of "book" (cf., Liebesman & Magidor, 2017). If the single-entry hypothesis is right, then the physical-token and abstract-content senses of "book" are available via the same lexical entry, so "blicket" will be mapped to that selfsame lexical entry and children should thus be immediately able to use it to express other senses of "book." If Fodor and Lepore are right, however, then mapping "blicket" to book-*qua*-token is insufficient to generate book-*qua*-content uses of "blicket."

Children then saw a story in which Ernie read a book which was physically long but short in content while Cookie Monster read a book which was physically short but long in content—though the words "book" and "blicket" were not used. In one condition, the books were read out loud, so the content sense of "book" was most relevant. Then, if Elmo summed up the story by saying "Ernie read the long blicket," the 4-year-olds judged him to be wrong. This shows that despite "blicket" being explicitly linked to the physical-token sense of "book," children freely and spontaneously extend the novel word to the abstract-content sense despite lacking general metalinguistic ability. Srinivasan and Snedeker (2011) also tested homonyms (e.g., mapping "davo" to "bat") and found that children did *not* spontaneously extend the word from one meaning (baseball bat) to another (flying bat). This result strongly suggests that polysemy arises out of modulation within a lexical entry rather than a sense of (e.g., phonological) relatedness across lexical entries.[10]

Srinivasan and Snedeker (2014) also found that young children will spontaneously extend *novel* words across polysemous senses even when the denotations are perceptually and taxonomically quite different. Once children are shown pictures of chickens as examples of "darpa," they will use "darpa" for grilled chicken meat rather than a duck, despite ducks being more visually/taxonomically similar to the original stimuli. Similarly, children who learn the word "buck" as a verb for some activity freely generate senses of "buck" as a noun that refers to the instrument used for that activity, even for completely novel activities/instruments (Srinivasan, Al-Mughairy, Foushee & Barner, 2017).

This evidence suggests that some generalizations about polysemy may reflect invariant conceptual structure rather than merely idiosyncratic linguistic convention. Supporting this conjecture, Srinivasan and Rabagliati (2015) found evidence for certain abstract polysemous transformations (e.g., container to containee, like "bottle") across 14 different languages (see also Zhu & Malt, 2014). Moreover, Srinivasan, Berner and Rabagliati (2019) showed that polysemy plays a crucial role in language acquisition. Famously, children shown an example of a "dax" will apply "dax" to novel objects that have a similar shape to the initial stimulus, showing a "shape bias" while ignoring other properties like color and size (Landau, Smith & Jones, 1988). Srinivasan et al. (2019) showed children and adults a material ("some *gup*") followed by an object made from that material ("a *gup*"). Subjects then saw an object with the very same shape but different material and an object with a completely different shape but the same material and were asked "Can you point to a *gup*?" Both children and adults selected the object with a different shape but the same material, showing that they understood the word to refer to both the material and objects made from that material. Thus abstract forms of polysemy are sufficiently important for language acquisition that they can trump robust aspects of language acquisition such as the shape bias.

[10]One might argue instead that polysemy is homonymy plus *actual relatedness among concepts* rather than a metalinguistic sense of relatedness (Devitt, Forthcoming). But this view fails to explain why homonyms *inhibit* each other as opposed to simply failing to prime each other (Maciejewski & Klepousniotou, 2020), or why children spontaneously extend novel words in ways characteristic of polysemy but *not* to visually or associatively related stimuli.

Rodd et al. (2002) estimate that 80% of words are polysemous. Even for technical terms with apparently rigid meanings like "neutron," we can freely generate new polysemous uses: "If a neutron was enlarged and put it in a woodchipper, there would be neutron all over the room!" While physically absurd, this sentence is semantically acceptable. This sort of spontaneous generation of count/mass noun polysemy is an example of the "universal grinder" (Pelletier, 1975)—a mechanism that transforms count nouns into mass nouns.

Polysemy is the rule rather than the exception. Its various forms express a fundamental, ubiquitous, and possibly innate structural feature of human lexicons.[11] And more to the point for present purposes, polysemous words involve a semantic representation that generates distinct denotations (see also Ortega Andrés & Vicente, 2019; Vicente, 2018). Given the common (and Fodorian) assumption that concepts are semantic representations, then concepts generate distinct denotations as well and Fodorian atomism is false.

4 | CONCEPTS AND WORD MEANINGS

4.1 | Semantic representations as nonconceptual pointers

A natural move on behalf of the Fodorian atomist would be to deny that semantic representations are concepts. At first glance, this move seems ad hoc (not to mention anathema to Fodor himself). Grasping word meanings is a core role for concepts, akin to their role in providing the vehicles of categorization. One could argue that *some* vehicles of categorization are not concepts, such as representations involved in color categorization (Block, n.d.). Similarly, one might argue that *some* word meanings are nonconceptually represented, such as complementizers like "that" or logical operators like "if" (Braine & O'Brien, 1991; cf., Peacocke, 1992). But this limited thesis amounts to a qualification on the general truth that concepts are semantic representations rather than a general denial of it. Our main question is not whether some meanings are nonconceptual, but rather whether meanings in general—meanings of ordinary lexical items such as "dog," "cake," and so forth—are represented conceptually.

Pietroski (2018) argues that word meanings are nonconceptual "instructions for how to fetch concepts" (p. 1). For Pietroski, distinct senses of a polysemous word are represented by distinct concepts, thus showing that the common word meaning is not itself conceptually represented (Pietroski, 2018, pp. 3ff; see also Glanzberg, 2011). Similarly, Recanati considers a view on which semantic representations have the "wrong format" for thought (Recanati, 2004, p. 140, Recanati, 2017). I propose to take it as a simple empirical question whether the vehicles

[11]Why should humans have an innate—or at least cross-culturally ubiquitous—apparatus to flexibly shift denotations of lexical items? Xu, Malt and Srinivasan (2017) speculate that the function of polysemy is to maintain a compact lexicon while facilitating a Humboldtian generation of infinite meanings by finite means (Chomsky, 1965). One aspect of this function may be to counterbalance the massive underdetermination of meaning in ordinary acquisition of concepts in childhood. Children often learn words at a single exposure and retain them for weeks (Carey & Bartlett, 1978). It is unclear how much information is represented in this "fast mapping," however; Carey (2010) writes, "[n]ever did anybody believe that children typically create full lexical representations upon just one or even a few exposures to a new word" (p. 184). Perhaps pre-existing schemata for generating novel polysemous senses prepare children systematically to pack new information into these sparse initial conceptions—alongside other pre-existing formal apparatuses, like "syntactic bootstrapping" (Gleitman, 1990). As Gleitman puts it, "semantics is much richer than syntax. But there's enough information in the syntax to point to the right neighborhood for the meaning of the verb. And then 'the world' has some hope of supplying the detail" (Gleitman et al., 2019, p. 11). The sorts of detail the world can supply may be enriched (and constrained) by generative procedures underlying polysemy.

we use for thinking about categories like *dog* also function as semantic representations (and these conceptual semantic representations could also function as instructions/pointers, a possibility pursued in Section 5). The hypothesis that the representations we deploy to grasp meanings are not the concepts we think with takes on some empirical commitments: (a) we should observe divergence between how semantic representations function and how concepts function, and (b) any interaction between language and cognition should require an *additional computational step* of translating between different representational formats.

4.2 | Evidence for conceptual semantic representations

Murphy (2002) argues that semantic representations are concepts by appeal to priming effects. Semantic representations appear to be organized in ways that mirror the taxonomy-based, typicality-based, and contiguity-based organization of concepts (Federmeier & Kutas, 1999). For example, processing "vehicle" is quicker if it is anaphorically dependent on "bus" as opposed to "tank," reflecting typicality and taxonomy in the lexicon (Garrod & Sanford, 1977; see Murphy, 2002, p. 396). Opponents may reply that the lexicon simply reduplicates these aspects of conceptual organization. Murphy (2002) judges this move to be ad hoc (p. 394). It is not ad hoc to think that nonconceptual lexical representations might be associatively linked, and that these associative links happen to mimic some aspects of conceptual organization. It would be ad hoc to reduplicate virtually all the organizational complexity of conceptual memory in the lexicon to preserve the hypothesis that semantic representations are nonconceptual. Nonetheless, it is helpful to look for independent evidence.

If semantic representations are concepts, then we should expect a tight interplay between semantic processing and other processes that use concepts, such as visual categorization. Indeed, words function as extremely effective cues to visual categorization and search (Boutonnet & Lupyan, 2015; Potter, 1975). Furthermore, visual categorization maps images to the very same representations used to retrieve the meanings of the corresponding words. Potter and Faulconer (1975) had subjects name basic-level images (e.g., a chair) and words ("chair"), or provide the superordinate category of the image or word ("furniture"). While naming a word was more than 200 ms quicker than naming the corresponding image—expected, since the ortho/phonological information is already deployed in reading—mapping an image of a chair to "furniture" takes *no longer* than mapping "chair" to "furniture." This result suggests that categorizing an image of a chair and retrieving the meaning of "chair" *both* require mapping stimuli to the concept CHAIR (and from there to FURNITURE), which is neither proprietarily linguistic nor perceptual.

Later, in addition to replicating this earlier result (finding a > 200 ms lag for naming images), Potter, Kroll, Yachzel, Carpenter and Sherman (1986) tested whether representations used for visual categorization can compose with semantic representations to represent sentence meanings. Subjects saw words presented serially to form a sentence like "Judy needed the stool to reach the lightbulb." In some conditions, however, one or two of the words (e.g., "stool" and "lightbulb") were replaced by images. Subjects were then asked to recall the sentence and evaluate it for plausibility. If semantic representations were nonconceptual, then performance should be slowed by at least 200 ms in the one-image condition and 400 ms in the two-image condition; however, they found that, depending on materials, performance was either equivalent or slowed by significantly less than 200 ms in the image conditions, and not significantly different between the one-image and two-image conditions. These results suggest that vehicles

of categorization can function *immediately* to grasp word meanings and compose with other semantic representations to grasp sentence meanings.

4.3 | Evidence for conceptual sentence-meaning representations

The evidence thus far suggests that semantic representations can function like concepts, and that concepts can function like semantic representations, including functioning compositionally in grasping sentence meanings. If these hypotheses are true, then we should predict that sentence-meaning representations function like full-blown thoughts. Sentence-meaning representations should, for example, function automatically, without an intermediating step, as premises in logical inferences. There is some evidence for this prediction.

Lea (1995) found that if subjects read sentences of the form "p" and "if p then q," they automatically perform the inference to "q" (tested by a lexical decision task for words semantically related to "q"). Interestingly, this was true even when performing the inference was irrelevant to forming a coherent understanding of the story. This suggests that sentence meanings are poised to be automatically inferentially promiscuous, just as propositional thoughts are (see also Rader & Sloutsky, 2002). Moreover, logical inference can be triggered through subliminal presentations of premises, again suggesting automaticity in logical inference from sentence meanings (Reverberi et al., 2013).

The automaticity and context-independence of these effects suggests that sentence-meaning representations function in inferences without having to be translated into a different format first. Furthermore, "p" and "if p then q" will only facilitate "q" if the premises occur near each other in the text (Lea et al., 2005). This result suggests that logical inferences are not merely run on "situation models" (Zwaan, 2016), that is, the postlinguistic representations we construct to make sense of described scenarios. Instead, the *initial* stages of sentence comprehension deploy representations useable in logical inferences. The simplest explanation is that sentence-meaning representations simply are thoughts.

4.4 | Polysemous thoughts

Another way into this dispute concerns polysemy resolution. For Pietroski, word meanings lack denotations, which are achieved through retrieving one of the concepts pointed to by the nonconceptual semantic representation. A driving assumption seems to be that concepts themselves are not polysemous. In that case we should not find cases where concepts are deployed but polysemy fails to be resolved—there must be no *polysemous thoughts*.

Though I know of no direct empirical evidence on this question, I think intuition tells against the prediction. Consider "door." One can say that John knocked on the door (*qua*-barrier) or that John walked through the door (*qua*-aperture). The meaning <door> is neutral between these denotations. Can we think with that neutral content?

Suppose you walk into a classroom and look for an open seat. I suggest that you can simply think THERE IS AN OPEN SEAT BY THE DOOR without resolving whether you have the barrier or aperture in mind.

Another polysemy case is "lunch," which can refer to the food one eats around noon ("Lunch was disgusting") or to a meeting ("That was a productive lunch"). Suppose you have lunch with an old friend. The food is delicious, the conversation is fun, and you leave feeling

happy about the whole affair. You might think LUNCH WAS GREAT without the token deployment of LUNCH referring to the meal or the meeting in particular.

Yet another case concerns countries: FRANCE can denote a piece of land ("France is hexagonal"), a government ("France is in the EU"), a population ("France is unhappy with its government"), a culture ("France produced their best films in the 1960s"), even a sports team ("France won the last World Cup"). You might tour France, become enamored with the landscape, the population, the culture, and so forth, and think I LOVE FRANCE without having a particular denotation in mind.[12]

Another set of examples include mental analogues of copredication sentences. Intuitively, somebody can think THIS BOOK HAS A RED COVER AND A BRILLIANT PREFACE, thus using BOOK in a way that allows it to denote book-*qua*-vehicle and book-*qua*-content. One could object that a correct structural description of the thought would take the form THIS $\text{BOOK}_{\text{VEHICLE}}$ HAS A RED COVER AND THIS $\text{BOOK}_{\text{CONTENT}}$ HAS A BRILLIANT PREFACE, with two token concepts of distinct types. But presumably one can think thoughts of the form X IS F AND G without iterating tokens of X.

Furthermore, some evidence from the concepts literature suggests that a single concept can have multiple denotations. Some concepts are "dual-character concepts" (Knobe, Prasada & Newman, 2013). SCIENTIST can denote somebody who fits the stereotype of a scientist (works in labs, etc.), or a "true scientist," who has an underlying drive for truth. Someone can be a true scientist even if they have never worked as a scientist, and someone might be a working scientist but fail to be a true scientist (Knobe et al., 2013). SCIENTIST can denote somebody who fits the stereotype of a scientist (works in la). Dual-character concepts look prima facie like cases where a concept yields distinct categorization judgments depending on whether one thinks with it in a typicality-based way, or an essence-based way (Newman & Knobe, 2019; see also Machery & Seppälä, 2011).

4.5 | Moving forward

We should resist the idea that semantic representations constitute a nonconceptual representational layer mediating language and thought. We should instead embrace the straightforward, classic idea that core elements of the mental representations of word meanings are simply concepts, perhaps accompanied by structural frames and other supplementary information.

Some readers may remain unconvinced and insist that semantic representations are nonconceptual. Even so, such readers should nonetheless find it independently interesting to explore what a theory of concepts would look like if we maintained the platitudinous thesis that concepts represent word meanings. In that case, polysemy successfully undermines Fodorian atomism about concepts. The next question is what to put in its place. One possibility, explored below, is to reject the traditional semantic component of atomism (i.e., that concepts fix denotations atomistically) while retaining the syntactic component (i.e., that concepts lack internal structure).

[12]This example opens up the possibility of nonpropositional attitudes featuring FRANCE that fail to disambiguate, such as liking France *tout court* (Grzankowski, 2016; Montague, 2007), perhaps by merely linking the concept with a positive valence.

5 | CONCEPTS AS POINTERS

5.1 | Pointer architectures

Pietroski and Glanzberg deny that semantic representations are conceptual, but they make an interesting positive claim: Semantic representations "point" to concepts or provide "instructions" to retrieve concepts. I now suggest a modification of this thesis: Semantic representations are conceptual, and concepts are atoms that point to information in long-term memory. A theory of concepts as pointers provides a simple, architecturally precise account of how concepts function in the mind. It also fills in a noted gap in atomistic theories, namely, explaining the relation between concepts and "conceptions," the bodies of information that fall under a concept in an individual mind.

Representations are stored at memory locations. Some representations contain *addresses*, that is, names of memory locations.[13] These representations are pointers—they "point" to the memory locations named in their addresses.[14] Gallistel and King (2010) show how pointer architectures can massively simplify computation. In some cases, it is useful to compute over a variable whose value changes frequently, such as the current position of the Sun. In that case, you may write the position of the Sun at a location in memory. Some computation that needs to make use of the Sun's location can then compute over a representation that addresses (points to) this location; the process may then be redirected to the pointed-to location where it may retrieve the symbol stored there that explicitly encodes the current location of the Sun.

Green and Quilty-Dunn (2017) argue that a pointer architecture underlies object files in working memory. For Gallistel and King (2010), pointer architectures are ubiquitous in computational systems due to "the ineluctable logic of physically realized computation" (p. 158). My aim here is not to defend this general thesis about computation. Instead, I propose that we understand *concepts* as pointers. Concepts are syntactically atomic representations that address memory locations. At those memory locations are stored a large array of representations, including prototypes, theories, and other informational structures. These informational structures constitute the "conception" that falls under the concept (Camp, 2015; Löhr, 2020; Rey, 1983). The concept itself is an unstructured symbol akin to a Fodorian atomistic concept.

Thinking of concepts in terms of a pointer architecture provides a simple, nonmetaphorical account of the relation between concepts and the bodies of information that fall under them. Importantly, one can compute over a pointer without computing over the information it points to—a form of "lazy" computation. Thus pointers can compose into discursive structures that function as premises in logical inference. Logical inferential rules can specify types of constituent structure and enable concepts to figure in logical inferences independently of the

[13]Memory locations are standardly called "addresses," but I use distinct terminology here to distinguish pointers from the locations they point to.

[14]This way of talking is in keeping with the notion of "location-based addressing" as opposed to "content-based addressing" (Frankland & Greene, 2019). An example of the latter is Eliasmith's (2013) and Blouw, Solodkin, Thagard & Eliasmith (2016) important notion of *semantic pointers*, which are compressions (e.g., statistical summaries) of sensorimotor information and "point" to that information in the sense that they compress it. In keeping with atomism, I will instead stick to location-based addressing in discussing concepts as pointers. The PET FISH problem arguably arises for semantic pointers (i.e., PET FISH need not compress PET or FISH-related sensorimotor information). I'm open to a role for compressions/summaries enabling "lazy" cognition that still carries useful information, but unlike Eliasmith (2013, p. 298), I think phrasal meanings like "pet fish" are grasped compositionally, in which case atomistic (and therefore location-based) pointers seem ineliminable.

information they point to (Quilty-Dunn & Mandelbaum, 2018; cf., Shea, n.d.). Concepts can also compose into complex concepts without composing their conceptions (such as prototypes), thereby avoiding PET FISH-type problems.

Understanding concepts as atomic pointers avoids the problem of messiness as well. What unifies disparate structures such as prototypes and theories under the same concept is not their tendency to be coactivated; it is the fact that they are stored together at a memory location pointed to by that concept.

The idea that concepts point to memory locations where conceptions are stored has some antecedents. For example, some have argued that concepts are akin to "labels" on mental "files," which contain assorted information (Fodor, 2008; Margolis, 1998; Recanati, 2012). The notion of a mental file is a useful construct (including in some experimental contexts, for example, Perner, Huemer & Leahy, 2015). Recanati (2012) cashes out the file metaphor in terms of "epistemically rewarding relations" between mental representations and their referents, which "enable the subject to gain information from the objects to which he stands in these relations" (p. 35). The idea that concepts have unique denotations lies at the core of the mental files theory, including the similar view of Margolis (1998) and Laurence and Margolis (1999) (and Millikan's, 2017 notion of "unicepts").[15] But since, as we have seen, concepts are polysemous, we cannot build inflexible denotations into the individuation conditions of concepts, nor count information as falling under a concept's conception by virtue of coreference (whether "de jure" [Fine, 2007] or otherwise). The idea of concepts as labels on files is problematic because it is a version of Fodorian atomism.

Instead, we should hold that concepts are pointers to memory locations.[16] There are no a priori restrictions on what representations can be stored at a memory location pointed to by a concept. It could include sentence-sized thoughts, term-sized predicates, nonconceptual sensorimotor images, mental models (Johnson-Laird, 2006), prototypes, theories, exemplars, ideals, "dual-character" representations (Knobe et al., 2013), and just about anything else. What unifies this congeries of conceptual structures is not any shared structural properties (such as representational format), semantic properties (such as coreference), epistemic properties (such as epistemically rewarding relations), or being *associated* with the pointer, which too many other things are as well. Instead, it is simply architectural: These representations are stored at the same memory location, and that memory location is addressed by the concept in question.[17] Talk of addressing memory locations is no more metaphorical in human memory than it is in other

[15]Recanati briefly suggests that mental files may be modulated in context in ways similar to the framework defended here (Recanati, 2012, p. 140). The thesis that concepts are generative pointers could be thought of as a version of the mental files theory as long as (a) files can regularly fail to determine reference and (b) the relation between files and their contents is pointing/addressing and not association or coreference.

[16]Talk of pointers in something like this sense is common in relevance theory (Carston, 2010; Sperber & Wilson, 1994). Like Glanzberg and Pietroski, relevance theorists often take pointers to point to concepts, not to be concepts themselves. Sperber and Wilson (1994), for example, describe a word meaning as "a pointer to a concept" (p. 196). Carston (2002) speculates that meanings are "not really full-fledged concepts, but rather concept schemas, or pointers to a conceptual space, on the basis of which ...an actual concept (an ingredient of a thought) is pragmatically inferred" (p. 360). These passages suggest a "wrong format" view (Carston, 2012; cf., Carston, 2019), unlike the view defended here. But the idea that semantic representations incorporate addresses of memory locations has been long defended by relevance theorists. In general, the theory defended in this paper is broadly congenial to relevance theory, *modulo* some concerns about ad hoc and metaphorical senses discussed below.

[17]There may be psychological laws constraining the information that can be housed under a concept due to general forms of polysemy. This possibility is explored below. What matters for present purposes is that the functional relation of concept to conception (pointing) places no constraints on the contents of conceptions.

computational systems, such as the biological ones described by Gallistel and King (2010) or the one I used to write these words. Functionally individuated memory locations house symbols; concepts point to those locations and thereby facilitate the retrieval of those symbols.

5.2 | Generative pointers

Fodorian atomists could make use of the notion of a pointer architecture along the lines sketched above as a way of fleshing out the "file" metaphor without tacit associationist assumptions. However, polysemy falsifies Fodorian atomism. This leads to the full positive view defended here: Concepts are not merely pointers; they are *generative pointers*.

To illustrate the core idea, consider the polysemous concept DOOR. DOOR is a pointer to a memory location where a large amount of information is stored pertaining to doors-*qua*-apertures and doors-*qua*-barriers. The concept/pointer itself fails to denote either and can figure in thought without resolving this ambiguity (as argued above). Upon deploying the concept, computational processes may be redirected to the addressed memory location. For simplicity's sake, suppose the location houses two visual images, one an exemplar of door-*qua*-barrier and one an exemplar of door-*qua*-aperture. Which image is then deployed—not merely activated, but deployed—modulates the denotation of DOOR.[18] When thinking JOHN PAINTED THE DOOR you retrieve, for example, an image of a brown slab with a doorknob on it, and thus your token thought denotes door-*qua*-barrier rather than door-*qua*-aperture. See also Figure 1, which provides a simplified diagram of CHICKEN as conceived by a generative pointer framework.

Denotation is secured not by concepts themselves but by functional interactions between concepts and bits of conceptions. Concepts are pointers that enable thinkers to generate denotations (and, therefore, truth-conditional propositional contents) through retrieving disambiguating information stored at the pointed-to locations.

Likely far more than mere sensorimotor imagery is involved in disambiguation. WATER points to an array of information, some of which is used to think about water-*qua*-natural-kind and some is used to think about water-*qua*-appearance-property. (The difference between these properties is roughly that the latter is, and the former is not, present on Twin Earth.) When we think about water-*qua*-natural-kind, we deploy WATER and exploit its pointing function to retrieve an essentialist theory that characterizes water (Gelman, 2003; Newman & Knobe, 2019; cf., Strevens, 2019). We could instead think about it as an appearance property by retrieving other information, resulting in judgments that swamp water is water but weak coffee is not, despite the latter having a higher H_2O content (Malt, 1994). The fact that these bodies of information sit side-by-side in the memory location pointed to by WATER explains why people prefer to say of Twin Earth water that it both is and is not water (Tobia, Newman & Knobe, 2020). There may even be a multiplicity of essentialist theories, some invoking causal essences like H_2O and others invoking teleological essences (Rose & Nichols, 2019), generating multiple kind denotations.

[18]It is unclear what the activation-deployment distinction amounts to. I suggest that three conditions are individually sufficient for use/deployment over mere activation: (a) figuring in some computational process (e.g., categorization); (b) being moved into working memory; or (c) surpassing some threshold of activation. (a) constitutes use in the most straightforward sense. There may be cases of idle thought where one deploys a representation without its figuring in any particular computation, as in (b) and (c). These forms of deployment may interact: Storage in working memory facilitates use in computational processes; being highly activated facilitates storage in working memory.

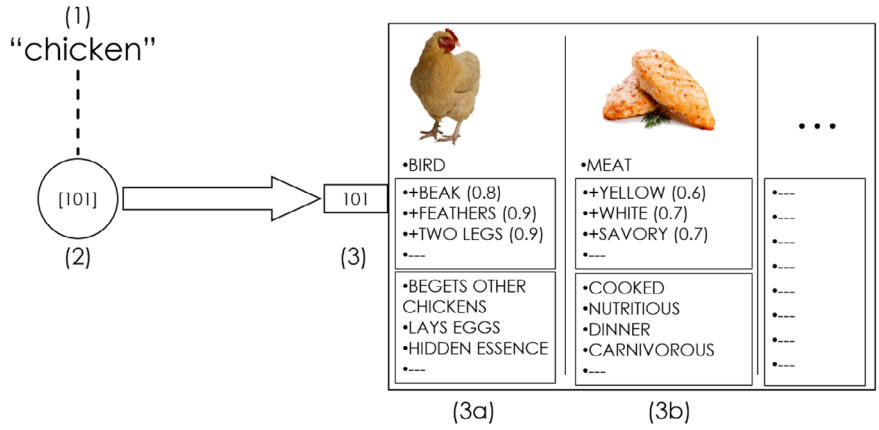

FIGURE 1 CHICKEN as a generative pointer. A lexical representation of the ortho-phonological properties of "chicken" (1) is conjoined in a lexical entry with the concept CHICKEN, an atomic symbol (2), which points to the memory location where the informational structures constituting the conception of CHICKEN are stored (3). Within the conception of CHICKEN, one chunk of information includes diverse structures pertaining to chicken-*qua*-bird (3a) and another includes diverse structures pertaining to chicken-*qua*-meat (3b), either of which can be retrieved to resolve polysemy [Color figure can be viewed at wileyonlinelibrary.com]

Polysemy resolution need not happen because pointed-to information descriptively specifies a denotation. Instead, there may be a nondescriptive reference relation, so long as the mental relatum of that relation is a functional interaction between some stored information and the concept/pointer.[19] It is not that the stored information provides a definite description, but merely that its retrieval via the concept/pointer functions to denote something. The vehicles of reference would then be events of retrieval via concepts. Perhaps we come to acquire an essentialist theory and store it under WATER because of appropriate reference-grounding interactions with water-*qua*-natural-kind. We can then retrieve the theory via WATER and arrive at that denotation in virtue of the historical connection even if the theory fails to descriptively pick out water-*qua*-natural-kind (perhaps because it posits a nonexistent telos—Rose & Nichols, 2019).

This paper began by stipulating that concepts are mental representations. Thus they must represent something (ignoring methodological solipsism; Fodor, 1980; Chomsky, 2000). I have argued that token events of retrieval via concepts represent something (e.g., doors-*qua*-barriers). But what do the concepts/pointers themselves represent? The answer to this question is not obvious.

A simple approach would be to hold that concepts represent arbitrary disjunctive properties consisting of the disjunction of denotations that can be fixed in polysemy resolution. Thus CHICKEN denotes <chicken-*qua*-bird or chicken-*qua*-meat or...>. It is quite counterintuitive to

[19]One possibility is that, as in Figure 1, concepts like CHICKEN contain atomic symbols like BIRD and MEAT which are retrieved to resolve polysemy. In that case, polysemy resolution may operate via variable binding (Gallistel & King, 2010), where the variable is CHICKEN and the possible values include BIRD and MEAT; the result of variable binding could then be denoted with subscripts: $\text{CHICKEN}_{\text{BIRD}}$ or $\text{CHICKEN}_{\text{MEAT}}$. This story need not require that concepts functioning as values like BIRD or MEAT have their own reference (on pain of regress—surely the concepts BIRD and MEAT are themselves polysemous). Instead, the result of variable binding, such as $\text{CHICKEN}_{\text{BIRD}}$, refers. This picture is appealingly simple, but I prefer to remain agnostic and make room for theories like Vicente's (2018), on which messy bodies of information figure in disambiguation.

think that our basic-level lexical concepts denote such arbitrary, unfamiliar properties. Moreover, if the point of a denotational view of conceptual content would be to provide compositional elements of truth-conditional contents, these disjunctions are ill-suited to do so. Insofar as a thought like THREE CHICKENS ARE FLYING AROUND THE BARN were to fix truth-conditions, they would involve the chicken-*qua*-bird sense of 'chicken' and *not* chicken-*qua*-meat.

Another possibility is that concepts do not denote at all. In a discussion of linguistic meaning, Harris (Forthcoming) argues that sentence meanings should be thought of not in terms of truth-conditional propositional contents, but in terms of *constraints* on which propositional contents can be recovered through pragmatic inference (Carston, 2012; Sperber & Wilson, 1995). For Harris, semantic constraints are evidential: Expressions constrain denotations and truth-conditional contents because they provide defeasible evidence about what speakers have in mind. Mental representations like concepts do not function as evidence in this way. But the notion of constraints on denotations could in principle be applied nonevidentially. Instead, perhaps concepts are constraints on *what can be denoted in a particular episode of thinking*. Perhaps the content of CHICKEN does not fix even a weird, disjunctive denotatum, but is instead a constraint on what can be mentally thought about via that concept.

If this sort of constraint semantics for concepts is correct, then unresolved polysemous thoughts like THERE IS AN OPEN SEAT BY THE DOOR do not have full-blooded propositional contents. Instead, its constituents provide constraints on which denotations can be achieved through them, and therefore on which propositional contents can be achieved through that type of thought. Despite not fixing a propositional content by itself, the thought THERE IS AN OPEN SEAT BY THE DOOR still serves core functions associated with propositional thought. For example, it can function as a premise in logical inferences and feed into linguistic processing independently of how the denotations of its constituents are resolved. This sort of independence allows us to explain what is invariant across thinking BOB HATES THE SCHOOL in the case where he thinks the administration makes terrible decisions and the case where he dislikes the architecture.

Where do these constraints come from? The deep answer to this question must lie in psychosemantics—that is, some story about the origin and function of concepts that makes sense of their semantic properties (Shea, 2018). But we can speculatively say a bit more. Many forms of polysemy are "regular" across domains and robustly cross-cultural. Srinivasan and Rabagliati's (2015) cross-cultural experiments found evidence that "polysemy is constrained by conceptual structure" (p. 124), which helps children "build a lexicon because learning one sense of a word could provide information about its other possible senses" (p. 148). The cross-cultural ubiquity of certain forms of polysemy suggests that the conceptions pointed to by our polysemous concepts may exhibit a structure that facilitates generativity in highly constrained, widely shared ways. Thus facts about how information is packaged—for instance, when a kind is identified, information about its material makeup and its function will be stored in separate packages at the same memory location—can play a role in explaining why certain referential successes are possible through a polysemous concept and others are not. To quote Fodor and Lepore (2002a): "Surely there just *couldn't be* a word that's polysemous between *lamb-the-animal* and (say) *beef-the-meat*? Or between *lamb-the-animal* and *succotash-the-mixed-vegetable*?" (p. 117). Such restrictions may arise due to structural features of our conceptions that enable rich conceptual generativity while imposing principled constraints on how concepts are used.

This (admittedly speculative) picture of constraints has the consequence that ad hoc modulations of word meaning, such as some metaphors, fall outside the constraints of a concept (Carston, 2012; Wilson & Carston, 2007). Independent priming data suggests metaphorical senses do not function like highly related polysemous senses, and are instead stored elsewhere

in the mind (Klepousniotou et al., 2008; Maciejewski, Rodd, Mon-Williams & Klepousniotou, 2020; cf., Floyd & Goldberg, 2020). Metaphorical senses and other ad hoc modulations appear to call on information beyond what is pointed to by the concepts deployed to grasp meaning—though perhaps conceptual content can change over time as one generation's metaphors transform into another generation's polysemous senses (Xu et al., 2017). The modulation of word meaning may extend beyond what concepts point to. In that case, it is not strictly true that a concept simply *is* a word meaning. Instead, a concept is a mental representation that is deployed to grasp word meanings, and a theory of conceptual content and a theory of word meanings are different enterprises entirely.

This possibility seems to accord with theoretical practice. For example, cognitive psychology is concerned with conceptual structure and content, but it does not seem thereby to provide a theory of lexical semantics. Nor does lexical semantics seem well-poised to provide a theory of concepts. Perhaps conceptual content involves constraints specified by the structure of information pointed to by concepts as suggested above, whereas word meanings are more open-ended, allowing for ad hoc and metaphorical modulation involving transitions across distinct concepts (Carston, 2019, Forthcoming). Another intriguing possibility is that, as Del Pinal (2018) forcefully argues, conceptual structure imposes principled semantic constraints on lexical modulation. But in either case, cross-linguistic and developmental evidence can give shape to a constraint semantics for concepts that accommodates polysemy while maintaining constraints on conceptual content.

6 | CONCLUSION

The foregoing has provided a sketch rather than a full theory of concepts. The core idea is that concepts are generative pointers. Specifically, concepts are atomic symbols that point to memory locations where conceptions are stored, and which bits of conceptions are retrieved through concepts determine the referent of a particular episode of thinking. This framework aims to unify concepts, conceptions, and representations of word meanings. Philosophical theories of concepts concerned with compositionality, psychological research on categorization and inference, and lexical semantics/pragmatics have all sat uneasily with each other for decades (Hampton, 2015). The hypothesis that concepts are generative pointers offers to bring them together. Lexical entries contain (inter alia) concepts, which are compositionally efficacious atomic mental representations that constrain denotations; concepts point to memory locations where rich informational structures are stored; these informational structures are available to guide inferences via the concept and to resolve the denotation achieved through the concept in thinking and interpreting utterances.

The flexibility of concepts on the generative pointer framework can be of substantial philosophical use. It is often assumed that certain concepts are "natural-kind concepts" and thus that they are used exclusively in a way that aims at denoting a hidden essence (Margolis, 1998) or some other basis of natural-kindhood (Boyd, 1991; Millikan, 1999; Strevens, 2019). A classic example would be WATER, which (suppose) aims at denoting an essence which turns out a posteriori to be H_2O (Kripke, 1980; Putnam, 1973). It is problematic for this view that we use WATER for categorization in ways that violate the natural-kind assumption: we judge tap water to be water and not Sprite, even though Sprite may have a higher H_2O percentage (Chomsky, 1995; Malt, 1994).

Does this entail that WATER is not a natural-kind concept? Or that natural-kind concepts are more minimalistic than essentialist theories (Strevens, 2019)? On the generative pointers framework, these questions are ill-posed. WATER does not have a unique denotation. It points to a memory location where diverse bodies of information are stored, including information about its appearance, its function in human life, and an essentialist theory (Gelman, 2003; Newman & Knobe, 2019). Part of an ordinary education may include placing the definition WATER IS H_2O at that location as well. On some occasions we may retrieve the essentialist theory, and/or the learned definition that cashes it out, and thereby think about water in a natural-kind way. On others we may retrieve information about its function and appearance, thereby yielding the judgment that tap water is water and Sprite is not. We may also rest content to categorize a candidate liquid as at once *water* and *not water* (Tobia et al., 2020).

This sort of shift may usefully apply to philosophical debates as well. For example, Byrne (2020, Section 2.6) argues that WOMAN denotes a biological kind by appeal (inter alia) to interchangeability of talk using "woman" and "female." This argument trades on the common assumption that FEMALE is a natural-kind concept *simpliciter*. Instead, like WATER, it may be polysemous and have salient non-natural-kind senses (Bettcher, 2013; See also Dembroff, 2020; Laskowski, 2020).[20]

Some philosophers argue for ameliorative modifications of concepts such as gender and race (Haslanger, 2000). The general practice of modifying concepts to suit desired ends is sometimes called "conceptual engineering" (Cappelen, 2018). We can understand conceptual engineering in a generative pointer architecture as involving (for example) the addition of a novel definition to the stock of information pointed to by a concept. This way of construing conceptual engineering accounts for the difference between merely "revising" a concept (i.e., adding new information to a conception) and changing concepts altogether (i.e., using a different conceptual pointer). The fact that the definition is stored together with so much other information also explains why it is so tempting to return to old cognitive habits. While a generative pointer architecture explains how conceptual engineering is possible, it also explains why it is difficult to maintain.

One major unresolved question concerns precisely how modulation works. Though this notion is crucial for explaining how we succeed in mentally securing denotations and truth conditions, I have not offered a theory of how we successfully generate a particular denotation. What determines which information we retrieve through CHICKEN on a particular occasion to generate a truth-evaluable thought? Until this question is answered, we have nothing more than a sketch of a generative theory of concepts. But to answer it we would have to determine how human beings decide what information to retrieve from long-term memory on different occasions and in different contexts (Sperber & Wilson, 1995). For that reason, it is possible that there is no theory forthcoming of how thinkers select some subset of information to modulate a pointer. To formulate such a theory would require diving headlong into the deepest, murkiest waters in cognitive science: domain-general, interest-sensitive decision processes in central cognition. The notion of generative pointers may provide some dim illumination in doing so.

ACKNOWLEDGEMENTS

Thanks to an anonymous reviewer for this journal, audiences at the Institute of Philosophy, New College of the Humanities, UCL, University of the Basque Country, and University of

[20]Byrne (2020, fn. 12) considers the possibility that "woman" is ambiguous (see also Byrne, n.d.), noting that it fails some ambiguity tests, but this is often true of polysemes (e.g., copredication may not be zeugmatic).

Warwick, participants in the Structure of Thought seminar at Oxford in 2018, and Robyn Carston, Alex Grzankowski, James Hampton, Dan Harris, Zoe Jenkin, Daniel Kodsi, Eric Mandelbaum, Matt Mandelkern, Michael Murez, Jesse Prinz, Joulia Smortchkova, and Richard Stillman for comments/discussion. Special thanks to Nick Shea for many helpful discussions, and to Agustín Vicente for helpful feedback and for interrupting my dogmatic Fodorian slumber.

ORCID


Jake Quilty-Dunn https://orcid.org/0000-0002-4642-8008

REFERENCES

Anderson, J. R. (1980). Concepts, propositions, and schemata: What are the cognitive units? *Nebraska Symposium on Motivation*, *28*, 121–162.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*, 261–295.

Apresjan, J. D. (1974). Regular polysemy. *Linguistics*, *12*(142), 5–32.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, *10*(1), 82–93.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(4), 629–654.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Emory symposia in cognition, 1: Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101–140). New York, NY: Cambridge University Press.

Bettcher, T. (2013). Trans women and the meaning of "woman". In A. Soble, N. Power & R. Halwani (Eds.), *Philosophy of sex: Contemporary readings* (6th ed., pp. 233–250). New York, NY: Rowman & Littlefield.

Bialystock, E. (1986). Factors in the growth of linguistic awareness. *Child Development*, *57*(2), 498–510.

Block, N. (n.d.). *The border between seeing and thinking*. Unpublished manuscript.

Blouw, P., Solodkin, E., Thagard, P. & Eliasmith, C. (2016). Concepts as semantic pointers: A framework and computational model. *Cognitive Science*, *40*, 1128–1162.

Boutonnet, B. & Lupyan, G. (2015). Words jump-start vision: A label advantage in object recognition. *Journal of Neuroscience*, *35*(25), 9329–9335.

Boyd, R. (1991). Realism, anti-foundationalism, and the enthusiasm for natural kinds. *Philosophical Studies*, *61* (1/2), 127–148.

Braine, M. D. & O'Brien, D. P. (1991). A theory of *if*: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, *98*(2), 182–203.

Brocher, A., Foraker, S. & Koenig, J. P. (2016). Processing of irregular polysemes in sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(11), 1798–1813.

Byrne, A. (2020). Are women adult human females? *Philosophical Studies*. https://doi.org/10.1007/s11098-019-01408-8

Byrne, A. (n.d.). *Gender muddle: Reply to Dembroff*. Unpublished manuscript. https://philpapers.org/rec/BYRG

Camp, E. (2015). Logical concepts and associative characterizations. In E. Margolis & S. Laurence (Eds.), *The Conceptual Mind: New Horizons* (pp. 591–621). Cambridge, MA: MIT Press.

Canal, P., Pesciarelli, F., Vespignani, F., Molinaro, N. & Cacciari, C. (2017). Basic composition and enriched integration in idiom processing: An EEG study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(6), 928–943.

Cappelen, H. (2018). *Fixing language: An essay on conceptual engineering*. Oxford: Oxford University Press.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Carey, S. (2010). Beyond fast mapping. *Language Learning, and Development*, *6*(3), 184–205.

Carey, S. & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference*, *15*, 17–29.

Carston, R. (2002). *Thoughts and utterances*. Oxford: Blackwell.

Carston, R. (2010). Metaphor: Ad hoc concepts, literal meaning and mental images. *Proceedings of the Aristotelian Society*, *110*, 295–321.
Carston, R. (2012). Word meaning and concept expressed. *Linguistic Review*, *29*(4), 607–623.
Carston, R. (2019). Ad hoc concepts, polysemy, and the lexicon. In K. Scott, B. Clark & R. Carston (Eds.), *Relevance, pragmatics and interpretation* (pp. 150–162). Cambridge: Cambridge University Press.
Carston, R. (2021). Polysemy: Pragmatics and sense conventions. *Mind & Language*, *36*(1), 108–133.
Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
Chomsky, N. (1995). Language and nature. *Mind*, *104*(413), 1–61.
Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.
Del Pinal, G. (2016). Prototypes as compositional constituents of concepts. *Synthese*, *193*, 2899–2927.
Del Pinal, G. (2018). Meaning, modulation, and context: A multidimensional semantics for truth-conditional pragmatics. *Linguistics and Philosophy*, *41*(2), 165–207.
Dembroff, R. (2020). Escaping the natural attitude about gender. *Philosophical Studies*. https://doi.org/10.1007/s11098-020-01468-1
Devitt, M. (Forthcoming). (2021). Semantic polysemy and psycholinguistics. *Mind & Language*, *36*(1), 134–157.
Eddington, C. M. & Tokowicz, N. (2015). How meaning similarity influences ambiguous word processing: The current state of the literature. *Psychonomic Bulletin & Review*, *22*(1), 13–37.
Eliasmith, C. (2013). *How to build a brain: A neural architecture for biological cognition*. Oxford: Oxford University Press.
Evans, G. (1982). *The varieties of reference*. Oxford: Oxford University Press.
Falkum, I. L. (2015). The *how* and *why* of polysemy: A pragmatic account. *Lingua*, *157*, 83–99.
Federmeier, K. D. & Kutas, M. (1999). A rose by any other name: long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*(4), 469–495. https://doi.org/10.1006/jmla.1999.2660
Fine, K. (2007). *Semantic relationalism*. Oxford: Blackwell.
Floyd, S. & Goldberg, A. (2020). Children make use of relationships across meanings in word learning. *Journal of Experimental Psycholoy: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000821
Fodor, J. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioral and Brain Sciences*, *3*(1), 63–73.
Fodor, J. (1998). *Concepts: Where cognitive science went wrong*. Oxford: Clarendon.
Fodor, J. (2003). *Hume variations*. Oxford: Clarendon.
Fodor, J. (2004). Having concepts: A brief refutation of the 20th century. *Mind & Language*, *19*(1), 29–47.
Fodor, J. (2008). *LOT2: The language of thought revisited*. Oxford: Clarendon.
Fodor, J. A. & Lepore, E. (2002a). The emptiness of the lexicon: Reflections on Pustejovsky. In *The compositionality papers* (pp. 89–119). Oxford: Oxford University Press.
Fodor, J. A. & Lepore, E. (2002b). The pet fish and the red herring: Why concepts still can't be prototypes. In *The compositionality papers* (pp. 27–42). Oxford: Oxford University Press.
Fodor, J. A. & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28* (1–2), 3–71. https://doi.org/10.1016/0010-0277(88)90031-5
Frankland, S. M. &, Greene, J. D. (2020). Concepts and compositionality: in search of the brain's language of thought. *Annual Review of Psychology*, *71*(1), 273–303. https://doi.org/10.1146/annurev-psych-122216-011829
Frazier, L. & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, *29*, 181–200.
Frisson, S. & Pickering, M. J. (1998). The processing of metonymy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1366–1383.
Gallistel, C. R. & King, A. P. (2010). *Memory and the computational brain: How cognitive science will transform neuroscience*. Oxford: Wiley-Blackwell.
Garrod, S. & Sanford, A. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior*, *16*(1), 77–90. https://doi.org/10.1016/s0022-5371(77)80009-1
Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford: Oxford Unversity Press.
Glanzberg, M. (2011). Meaning, concepts, and the lexicon. *Croatian Journal of Philosophy*, *11*(31), 1–29.

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*(1), 3–55.
Gleitman, L. R., Liberman, M. Y., McLemore, C. A. & Partee, B. H. (2019). The impossibility of language acquisition (and how they do it). *Annual Review of Linguistics*, *5*, 1–24.
Green, E. J. & Quilty-Dunn, J. (2017). What is an object file? *British Journal for the Philosophy of Science*. https://doi.org/10.1093/bjps/axx055
Grzankowski, A. (2016). Attitudes towards objects. *Noûs*, *50*(2), 314–328.
Hampton, J. A. (1988). Overextension of conjunctive concepts: Evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(1), 12–32.
Hampton, J. A. (1991). The combination of prototype concepts. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 91–116). Hillsdale, NJ: Lawrence Erlbaum Associates.
Hampton, J. A. (2000). Concepts and prototypes. *Mind & Language*, *15*(2–3), 299–307.
Hampton, J. A. (2015). Concepts in the semantic triangle. In E. Margolis & S. Laurence (Eds.), *The conceptual mind: New directions in the study of concepts* (pp. 655–676). Cambridge, MA: MIT Press.
Hampton, J. A. & Jönsson, M. L. (2012). Typicality and compositionality: The logic of combining vague concepts. In M. Werning, E. Machery & W. Hinzen (Eds.), *The Oxford handbook of compositionality* (pp. 385–402). New York, NY: Oxford University Press.
Harris, D. W. (Forthcoming). Semantics without semantic content. *Mind & Language*.
Haslanger, S. (2000). Gender and race: (What) are they? (What) do we want them to be? *Noûs*, *34*(1), 31–55.
Heim, I. R. (1982). *The semantics of definite and indefinite noun phrases*. (Doctoral dissertation). University of Massachusetts, Amherst.
Jackendoff, R. S. (1992). *Languages of the mind: Essays on mental representation*. Cambridge, MA: MIT Press.
Johnson-Laird, P. (2006). *How we reason*. Oxford: Oxford University Press.
Kaplan, D. (1978). On the logic of demonstratives. *Journal of Philosophical Logic*, *8*, 81–98.
Kawamoto, A. H., Farrar, W. T. & Kello, C. T. (1994). When two meanings are better than one: Modeling the ambiguity advantage using a recurrent distributed network. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1233–1247.
Klein, D. E. & Murphy, G. L. (2001). The representation of polysemous words. *Journal of Memory and Language*, *45*, 259–282.
Klepousniotou, E. & Baum, S. R. (2007). Disambiguating the ambiguity effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics*, *20*, 1–24.
Klepousniotou, E., Pike, G. B., Steinhauer, K. & Gracco, V. (2012). Not all ambiguous words are created equal: An EEG investigation of homonymy and polysemy. *Brain and Language*, *123*, 11–21.
Klepousniotou, E., Titone, D. & Romero, C. (2008). Making sense of word senses: The comprehension of polysemy depends on sense overlap. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1534–1543.
Knobe, J., Prasada, S. & Newman, G. (2013). Dual character concepts and the normative dimension of conceptual representation. *Cognition*, *127*, 242–257.
Kripke, S. (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.
Kutas, M. & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential. *Annual Review of Psychology*, *62*, 621–647.
Kutas, M. & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.
Landau, B., Smith, L. B. & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*, 299–321.
Laskowski, N. G. (2020). Moral constraints on gender concepts. *Ethical Theory and Moral Practice*, *23*, 39–51. https://doi.org/10.1007/s10677-020-10060-9
Laurence, S. & Margolis, E. (1999). Concepts and cognitive science. In *Concepts: Core readings* (pp. 3–81). Cambridge, MA: MIT Press.
Lea, R. B. (1995). On-line evidence for elaborative logical inferences in text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(6), 1469–1482.
Lea, R. B., Mulligan, E. J. & Walton, J. L. (2005). Accessing distant premise information: How memory feeds reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(3), 387–395. https://doi.org/10.1037/0278-7393.31.3.387

Liebesman, D. & Magidor, O. (2017). Copredication and property inheritance. *Philosophical Issues*, *27*(1), 131–166.

Löhr, G. (2020). Concepts and categorization: Do philosophers and psychologists theorize about different things? *Synthese*, *197*, 2171–2191.

MacGregor, L. J., Bouwsema, J. & Klepousniotou, E. (2015). Sustained meaning activation for polysemous but not homonymous words: Evidence from EEG. *Neuropsychologia*, *68*, 126–138.

Machery, E. (2009). *Doing without concepts*. New York, NY: Oxford University Press.

Machery, E. & Seppälä, S. (2011). Against hybrid theories of concepts. *Anthropology and Philosophy*, *10*, 99–126.

Maciejewski, G. & Klepousniotou, E. (2020). Disambiguating the ambiguity advantage effect: Behavioral and electrophysiological evidence for semantic competition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0000842

Maciejewski, G., Rodd, J. M., Mon-Williams, M. & Klepousniotou, E. (2020). The cost of learning new meanings for familiar words. *Language, Cognition, & Neuroscience*, *35*(2), 188–210.

Malt, B. C. (1994). Water is not H_2O. *Cognitive Psychology*, *27*, 41–70.

Marcel, A. J. (1980). Conscious and preconscious recognition of polysemous words: Locating the selective effects of prior verbal context. In R. S. Nickerson (Ed.), *Attention and performance VIII* (pp. 435–457). Hillsdale, NJ: Erlbaum.

Margolis, E. (1998). How to acquire a concept. *Mind & Language*, *13*(3), 347–369.

McNamara, T. P. (1992). Theories of priming: I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(6), 1173–1190.

Medin, D. L. & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*(3), 207–238.

Millikan, R. G. (1999). Historical kinds and the "special sciences". *Philosophical Studies*, *95*(1–2), 45–65.

Millikan, R. G. (2017). *Beyond concepts: Unicepts, language, and natural information*. New York, NY: Oxford University Press.

Montague, M. (2007). Against propositionalism. *Noûs*, *41*(3), 503–518.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Murphy, G. L. (2016). Is there an exemplar theory of concepts? *Psychonomic Bulletin & Review*, *23*, 1035–1042.

Newman, G. E. & Knobe, J. (2019). The essence of essentialism. *Mind & Language*, *34*(5), 585–605.

Nunberg, G. (1979). The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy*, *3*(2), 143–184.

Nunberg, G., Sag, I. A. & Wasow, T. (1994). Idioms. *Language*, *70*(3), 491–538.

Ortega Andrés, M. & Vicente, A. (2019). Polysemy and co-predication. *Glossa: A Journal of General Linguistics*, *4* (1), 1–23.

Peacocke, C. (1992). *A study of concepts*. Cambridge, MA: MIT Press.

Pelletier, F. J. (1975). Non-singular reference: Some preliminaries. *Philosophia*, *5*(4), 451–465.

Perner, J., Huemer, M. & Leahy, B. (2015). Mental files and belief: A cognitive theory of how children represent belief and its intensionality. *Cognition*, *145*, 77–88.

Peterson, R. R., Burgess, C., Dell, G. S. & Eberhard, K. M. (2001). Dissociation between syntactic and semantic processing during idiom comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1223–1237.

Pietroski, P. (2018). *Conjoining meanings: Semantics without truth values*. New York, NY: Oxford University Press.

Potter, M. C. (1975). Meaning in visual search. *Science*, *187*(4180), 965–966.

Potter, M. C. & Faulconer, B. A. (1975). Time to understand pictures and words. *Nature*, *253*, 437–438.

Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E. & Sherman, J. (1986). Pictures in sentences: Understanding without words. *Journal of Experimental Psychology: General*, *115*(3), 281–294.

Prinz, J. J. (2002). *Furnishing the mind: Concepts and their perceptual basis*. Cambridge, MA: MIT Press.

Prinz, J. J. (2011). Has mentalese earned its keep? On Jerry Fodor's *LOT2*. *Mind*, *120*(478), 485–501.

Prinz, J. J. (2012). Regaining composure: A defence of prototype compositionality. In M. Werning, E. Machery & W. Hinzen (Eds.), *The Oxford handbook of compositionality* (pp. 437–453). New York, NY: Oxford University Press.

Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.

Putnam, H. (1973). Meaning and reference. *The Journal of Philosophy*, *70*(19), 699–711.

Quilty-Dunn, J. & Mandelbaum, E. (2018). Inferential transitions. *Australasian Journal of Philosophy*, *96*(3), 532–547.

Rader, A. W. & Sloutsky, V. M. (2002). Processing of logically valid and logically invalid conditional inferences in discourse comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 59–68.

Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.

Recanati, F. (2012). *Mental files*. Oxford: Oxford University Press.

Recanati, F. (2017). Contextualism and polysemy. *Dialectica*, *71*(3), 379–397.

Reverberi, C., Pischedda, D., Burigo, M. & Cherubini, P. (2012). Deduction without awareness. *Acta Psychologica*, *139*(1), 244–253. https://doi.org/10.1016/j.actpsy.2011.09.011

Rey, G. (1983). Concepts and stereotypes. *Cognition*, *15*(1–3), 237–262. https://doi.org/10.1016/0010-0277(83)90044-6

Rodd, J. M., Cai, Z. G., Betts, H. N., Hanby, B., Hutchinson, C. & Adler, A. (2016). The impact of recent and long-term experience on access to word meanings: Evidence from large-scale internet-based experiments. *Journal of Memory and Language*, *87*, 16–37.

Rodd, J. M., Gaskell, G. & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, *46*, 245–266.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 28–49). Hillsdale, NJ: Erlbaum.

Rose, D. & Nichols, S. (2019). Teleological essentialism. *Cognitive Science*, *43*(4), e12725.

Rubenstein, H., Garfield, L. & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, *9*, 487–494.

Shea, N. (2018). *Representation in cognitive science*. Oxford: Oxford University Press.

Shea, N. (n.d.). *Drawing on the content of a concept*. Unpublished manuscript.

Shen, W. & Li, X. (2016). Processing and representation of ambiguous words in Chinese reading: Evidence from eye movements. *Frontiers in Psychology*, *7*(1713), 1–11.

Smith, E. E. & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Smith, E. E. & Osherson, D. N. (1984). Conceptual combination with prototype concepts. *Cognitive Science*, *8*, 337–361.

Sperber, D. & Wilson, D. (1994). The mapping between the mental and public lexicon. In P. Carruthers & J. Boucher (Eds.), *Language and thought: Interdisciplinary themes* (pp. 184–200). Cambridge: Cambridge University Press.

Sperber, D. & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford: Blackwell.

Sperber, R. D., McCauley, C., Ragain, R. D. & Weil, C. M. (1979). Semantic priming effects on picture and word processing. *Memory & Cognition*, *7*(5), 339–345. https://doi.org/10.3758/bf03196937

Srinivasan, M., Al-Mughairy, S., Foushee, R. & Barner, D. (2017). Learning language from within: Children use semantic generalizations to infer word meanings. *Cognition*, *159*, 11–24.

Srinivasan, M., Berner, C. & Rabagliati, H. (2019). Children use polysemy to structure new word meanings. *Journal of Experimental Psychology: General*, *148*(5), 926–942.

Srinivasan, M. & Rabagliati, H. (2015). How concepts and conventions structure the lexicon: Cross-linguistic evidence from polysemy. *Lingua*, *157*, 124–152.

Srinivasan, M. & Snedeker, J. (2011). Judging a *book* by its cover and its contents: The representation of polysemous and homophonous meanings in four-year-old children. *Cognitive Psychology*, *62*, 245–272.

Srinivasan, M. & Snedeker, J. (2014). Polysemy and the taxonomic constraint: Children's representation of words that label multiple kinds. *Language Learning and Development*, *10*(2), 97–128.

Stalnaker, R. (2014). *Context*. Oxford: Oxford University Press.

Stolz, J. A., Besner, D. & Carr, T. H. (2005). Implications of measures of reliability for theories of priming: Activity in semantic memory is inherently noisy and uncoordinated. *Visual Cognition*, *12*(2), 284–336.

Strevens, M. (2019). *Thinking off your feet: How empirical psychology vindicates armchair philosophy*. Cambridge, MA: Harvard University Press.

Titone, D. A. & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of Pragmatics*, *31*(12), 1655–1674. https://doi.org/10.1016/s0378-2166(99)00008-9

Tobia, K. P., Newman, G. E. & Knobe, J. (2020). Water is and is not H2O. *Mind & Language*, *35*(2), 183–208.
Van den Bussche, E., Notebaert, K. & Reynvoet, B. (2009). Masked primes can be genuinely semantically processed: A picture prime study. *Experimental Psychology*, *56*(5), 295–300.
Vicente, A. (2018). Polysemy and word meaning: An account of lexical meaning for different kinds of content words. *Philosophical Studies*, *175*(4), 947–968.
Vicente, A. & Martínez Manrique, F. (2016). The big concepts paper: A defence of hybridism. *British Journal for the Philosophy of Science*, *67*, 59–88.
Weiskopf, D. A. (2009a). Atomism, pluralism, and conceptual content. *Philosophy and Phenomenological Research*, *79*(1), 131–163.
Weiskopf, D. A. (2009b). The plurality of concepts. *Synthese*, *169*, 145–173.
Wilson, D. & Carston, R. (2007). A unitary approach to lexical pragmatics: Relevance, inference, and ad hoc concepts. In N. Burton-Roberts (Ed.), *Advances in pragmatics* (pp. 230–260). Basingstoke: Palgrave Macmillan.
Xu, Y., Malt, B. C. & Srinivasan, M. (2017). Evolution of word meanings through metaphorical mapping: Systematicity over the past millennium. *Cognitive Psychology*, *96*, 41–53.
Yolton, J. (1984). *Perceptual acquaintance: From Descartes to Reid*. Minneapolis: University of Minnesota Press.
Zhu, H. & Malt, B. C. (2014). Cross-linguistic evidence for cognitive foundations of polysemy. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *36*, 934–939.
Zwaan, R. A. (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin & Review*, *23*(4), 1028–1034.

How to cite this article: Quilty-Dunn J. Polysemy and thought: Toward a generative theory of concepts. *Mind Lang*. 2021;36:158–185. https://doi.org/10.1111/mila.12328