INTRODUCTION

In recent years, there has been an increasing interest in questioning the prospects of a unified theory of concepts, the scientific relevance of ‘concept’ as a theoretical term, and, more generally, the contribution of philosophy to the scientific study of concepts. Criticisms are not entirely unjust, since even the most popular accounts of concepts have failed to surmount problems that the community of cognitive scientists consider explanatorily important. One of the most notorious challenges faced by most researchers interested in the naturalistic study of concepts is to provide a convincing, albeit provisional, answer to the question of what concepts are. I will refer to this challenge as the ontological question regarding concepts. This task is often bracketed in favor of programmatic conceptions of the notion of a concept that leaves preliminary assumptions about the nature of concepts unexplained. To show that this is the case, I distinguish between epistemically and metaphysically motivated notions of a concept (EpiC and MetC, respectively) and consider the insurmountable problems of predominant models of EpiC to establish the conditions that are metaphysically constitutive of the nature of concepts. These kinds of concerns pertain to the ontological analysis on which MetC hangs and I intend to bring them to the fore as an attempt to defend the contribution of non-experimental philosophy to the empirical study of concepts.

It should be clear from the outset that I will be working within the framework of a representational theory of mind (RTM) that appeals to mental states as relations to mental representations with both semantic and causal properties to explain the behavior of creatures like us (Fodor, 1994). In the species of RTM that I endorse, thinking is understood in terms of mental processes enabled by mental representations with a constituent structure. A common assumption underlying representational theories of concepts in cognitive science is that complex mental structures
are made up of less complex constituents. Accordingly, concepts have been identified with those mental representations deemed to be the basic units of thought (Prinz, 2002).¹ The answers to the ontological question offered by MetC and EpiC are to be considered within this general framework. I will therefore keep the scope of the current work to theories of concepts that assume RTM to be their natural home.

While this initial demarcation serves the purpose of this article, some justification is in order, as several proposals about concepts stem from criticisms aimed at the classical paradigm of cognition in cognitive science.² This is the case of neo-empiricist views about concepts within the field of so-called embodied cognition (e.g., Barsalou 1999; Gallese & Lakoff, 2005; Prinz and Clark, 2004). Among the reasons against these alternative views is that the available evidence falls short of establishing the type of format (e.g., sensorimotor) and processing (e.g., simulations of perceptual and motor states) they attribute to concepts (Pylyshyn, 2002; Mahon & Caramazza, 2008; Machery, 2009; Dove, 2009). Another reason has to do with the dependence of a theory of concepts on a background theory that acts as a broader foundation for a viable cognitive science (Rescorla, 2020). While no detailed defense of a computational/representational theory of mind (henceforth, CRTM) will be deployed in this paper, it is worth noting that, unlike the aforementioned alternative approaches, the mental representation view of a concept finds support from CRTM as a background theory. A version of this theory offers a plausible causal explanation of the cognitive mind in terms of mental representations and processes where concepts play a fundamental role (Cain, 2002). This contrasts sharply with alternative views which are (explicitly or implicitly) committed to the metaphysical notion

¹ In this paper, I use the term “thought” as an umbrella term for cognitive states with constituent structure.

² I would like to thank an anonymous referee for raising this point.

of ‘embodied experience,’ where the very term of ‘embodiment’ is often used inconsistently among its advocates (Rohrer, 2007). I shall therefore adopt the working assumption that the viability of a representational theory of concepts is tied to the success of CRTM as the best available foundation for a scientific account of the mind.

The claim that concepts are structural constituents of thoughts is relatively uncontroversial among theorists who identify concepts with mental representations, but the widely shared view that concepts are themselves structured entities has raised a great deal of disagreement (Laurence & Margolis, 2022). My task in this paper will be to show that such disagreement is based on the unwarranted assumption that primitive concepts (i.e., the primitive stock of concepts out of which thoughts are constructed) must have internal structure if they are to account for the explanatory roles that cognitive scientists have attributed to what they call prototypes, exemplars, and theories. Machery (2009, 2010) has agreeably argued that none of the leading theories identifying concepts with any one of these constructs has managed to explain all the known phenomena related to work on, e.g., inference, categorization, and reasoning. In turn, Fodor (1998) has criticized what he claims is a dominant methodological doctrine in cognitive science, namely, that concepts are invariably identified with certain cognitive capacities. Pace orthodoxy in cognitive science, I will maintain that an argument for the explanatory effectiveness of prototypes, exemplars, or theories is orthogonal to an argument for the structure of primitive concepts.

Heterogeneous data figuring in experimental tasks have motivated the development of empirical theories positing dissimilar paradigms of conceptual structure (Murphy, 2002). Moreover, the failure to organize all the available findings into one coherent theoretical framework has encouraged the rise of pluralist and eliminativist positions regarding concepts. For instance, Laurence & Margolis (1999) have defended the view that concepts must have
a variety of structures, each of them responsible for the explanation of different phenomena. In turn, Machery (2009, 2010, 2015) has specifically advocated the elimination of the term ‘concept’ from scientific taxonomy. His proposed alternative to concepts is centered on the existence of several heterogeneous and independent psychological kinds that he characterizes in terms of default bodies of knowledge (BoKs). However, the inference from empirical evidence of heterogeneous data regarding what we know about x’s to the conclusion that we represent things in our mind in virtue of characteristically structured psychological kinds is fallacious. Important and deflationary reinterpretation of the notion of BoKs will be recommended.

The present paper is organized in the following way. In section 2, I address the notion of mental representations in cognitive science and the centrality of concepts in a computational and representational theory of mind. In section 3, I put forward and illustrate the distinction between EpiC (subsection 3.1) and MetC (subsection 3.2) relevant to the ontological question. Finally, in section 4, I present some concluding remarks, stressing the need for a reinterpretation of what cognitive scientists generally call prototypes, exemplars, and theories.

1. Mental representations and the centrality of concepts

The notion of representation is a fundamental construct in cognitive science, where thinking has been traditionally understood in terms of representational structures and computational procedures operating on those structures (Thagard, 2023). In this conception of how thinking works, concepts occupy a central place. To fully appraise the centrality of concepts, let us briefly unpack the notions of representation and computational procedures.

Following Von Eckardt (1993, 2003), mainstream cognitive science is committed to two basic working assumptions about the human cognitive mind. One is that the mind is a computational device. The other is that the mind is a representational or symbolic device. Certain explanatory roles that motivate the appeal to mental representations in the field:

First, the positing of mental representations is supposed to explain the intentionality of our cognitive capacities, as well as other properties such as their coherence and reliability that presuppose that intentionality. Second, most cognitive scientists also believe that the positing of mental representations is required to explain the kind of processing involved in the exercise of our capacities, namely, information processing, even when that processing is unconscious. (Von Eckardt, 2003, p. 428)

In line with the first assumption, mental processes are characterized as transformational procedures operating over sequences of token mental representations. The key idea here is that mental activity involves Turing-style computation over syntactically structured representations (Rescorla, 2020). With this picture in place, proponents of a CRTM have sought to answer the question of how rationality is mechanically possible (Rescorla, 2019).

For clarity, it helps to consider the intuitive notion of mental states expressed in terms of propositional attitudes. We normally explain other people’s behavior by appealing to their mental states (e.g., Emma did q because she wanted p and believed that not-p unless q). What this type of intentional explanation suggests is that reason-giving explanations are a species of bona fide causal explanations (Bermúdez, 2005, chapter 3). Within naturalistic theorizing, the question about the mechanical implementation of rationality reduces to providing a causal explanation at a physical level that vindicates the causal work that intentional explanations do. Thus, the account offered by CRTM according to which causal transitions between states result from formal, structural relations between these states becomes a plausible option. In this account, reason-giving relations at an intentional level are allowed to be mirrored in causal transitions at the level of formal syntactic properties of semantically evaluable representations (Burwood et al., 2003). This latter idea that our mental processes can be physically
Primitive Concepts and the Ontological Question

Bernardo Antonio Pino Rojas

implemented as interpreted formal symbols (i.e., information-baring structures) takes us to the second basic working assumption in mainstream cognitive science.

The view that the mind is a representational (or symbolic) device supplements the computational understanding of mental processes. Paradigmatically, cognitive scientists have conceived of the notion of a mental representation as a mental sentence type (i.e., a general sort of thing) that expresses a proposition (Vallejos, 2008). Based on this conception, mental states can be characterized as relations between thinkers and tokens of mental representations (i.e., particular instances) expressing specific propositional contents. As explained above, it is the formal properties of mental representations that determine how they are manipulated computationally. However, this should not lead one to believe that these formal properties are essential or constitutive properties of our mental representations. What is essential to the notion of a mental representation is that it has semantic properties, which makes it fit for a plausible explanation of the intentionality of our cognitive capacities.

Concepts are the least complex constituents of cognitive states, notably those states with constituent structures involved in the kind of mental processes that underwrite human cognitive capacities. Accordingly, concepts are deemed to play a central role in practically any theory that appeals to mental representations with semantic and causal properties to explain the behavior of creatures like us. The reason for this is that it is the semantic and causal properties of the constituent concepts of our mental states that determine the properties by which these states can explain, on the one hand, the intentional and causal properties of the organism (or different organisms) can be related in two characteristic ways to tokens of the same mental representation which means that a vaccine for COVID-19 will end the pandemic. In contrast, to believe that the Earth is flat involve the same kind of relation to two different mental representations. Hence, the claim that, within the framework of a representational theory of mind, a mental representation is individuated or distinguished by its intentional content.

Four prominent theories that identify concepts with mental representations stand out. Following Laurence & Margolis (1999), the Prototype theory holds that concepts have a statistical structure encoding information about properties that members of a category tend to possess. The Exemplar theory tends to identify concepts with sets of representations of category instances, also called exemplars (e.g., Medin & Schaffer, 1978; Nosofsky, 1988). In turn, advocates of the Theory theory of concepts (e.g., Carey, 1985; Keil, 1989; Gopnik et al., 2004) defend the general view that cognition resembles scientific reasoning. Here, concepts

---

3 For a relevant discussion on theoretical terms and concepts in science, see, e.g., Bunge (2017).
are variably identified with theoretical terms and mini theories of a given category (Prinz, 2002). Conceptual atomism is the view that concepts are mental representations with no structure (Fodor, 1998). Unfortunately, while each theory can be said to have its strengths comparatively, all of them have been subject to important criticisms. For reasons of space, suffice it to mention key challenges to each of them:4

Because the prototype of a complex concept (say, TROPHY HEAD) can hardly derive all its properties from the prototypes of its constituents (say, TROPHY and HEAD), the Prototype theory fails to provide an account of compositionality (i.e., conceptual combination). Neither the Exemplar theory nor the Theory theory can explain compositionality due to the same type of limitation. Similarly, it is also a serious problem for the corresponding theory that people can have concepts for which they do not have a prototype (e.g., SHIRTS THAT WERE BOUGHT IN JUNE), an exemplar (e.g., HOUR), or a correct theory (e.g., WHALE when whales are believed to be a type of fish). Additionally, none of these theories can provide an adequate account of intentionality (i.e., how the reference of a concept is determined), or publicity (i.e., that concepts can be shared by different people or the same individual at different times). Conceptual atomism has been criticized for being explanatorily impotent, unable to distinguish both coextensive and empty concepts, and radically committed to innate concepts (see section 3.2.1).

Inevitably, studying concepts empirically commits the researcher to certain preliminary assumptions about their nature (Millikan, 1998). If this is the case, that is, if empirical theories of concepts are implicitly or explicitly committed to certain ontological assumptions, then choosing between competing theories will have to consider the answer that those theories provide to the question of what concepts are. So, I will take a detailed account of the usual inventory of criticisms for granted and focus on what I deem to be one the most serious problems in the study of concepts, namely, the mistaken presumption that, insofar as the theoretical notion of a concept is intended to play a central explanatory role in a science of the mind, only complex concepts are worth our attention when assessing competing theories of concept individuation.5

2. EpiC and MetC kinds: Two doctrines of what Concepts are

Authors sometimes defend conceptions of natural kinds which, in their view, are relevant to philosophy of science but irrelevant to metaphysics (Brigandt, 2003; Griffiths, 2004). A similar situation arises in the literature on concepts, where some researchers (e.g., Machery, 2009; Piccinini & Scott, 2006) dismiss certain theoretical models of concepts which presumably fail to designate genuinely psychological kinds because they do not meet certain methodological constraints, while others (e.g., Fodor, 1998, 2004; Rey, 1983, 1985) argue against theories of concepts that disregard the implications for cognitive science of the metaphysical properties implicitly attributed to the constructs they posit. Thus, two broad positions can be distinguished. According to one, inquiry into the nature of concepts is exhausted by methodological demands. According to the other, it is not.

It is common among cognitive scientists to account for the possession of concepts in terms of individuals’ cognitive capacities. A natural reason for this is that cognitive abilities are amenable to

4 General reviews of the leading theories in philosophy and experimental cognitive psychology can be found in Laurence & Margolis (1999), Prinz (2002, chap. 3–4), Machery (2009, Chapter 4), and Murphy (2002).

5 I take the terms “simple concept” and “unstructured concept” to be coextensive. By contrast some authors (e.g., Smith & Osherson, 1984) have addressed the notion of a simple concept within the framework of the structured model of a concept they defend (e.g., simple vs complex prototypes).
experimental study, which is another way of saying that they can be specified in terms of psychological phenomenon (see 3.1.1). To the extent that this practical constraint is motivated by the role of concept possession in conceptual identity, it is usually assumed as a methodological requirement for concept individuation. Indeed, as some have argued (e.g., Edwards, 2009; Fodor, 1998), possession conditions for concepts are (explicitly or implicitly) taken to be ontologically important to the kind of things concepts are by most psychological theories of concepts. When a theory of concepts is committed to this latter assumption, an epistemically motivated doctrine to the ontological question will be obtained. Otherwise, a metaphysically motivated doctrine will be obtained. Based on these two doctrines, I will distinguish between an epistemically motivated notion of a concept (EpiC) and a metaphysically motivated notion of a concept (MetC), respectively.

The opposition between EpiC and MetC should not be taken to imply a deprecating attitude towards empirical inquiry about the mind on the part of theorists adopting MetC. Indeed, both approaches are compatible with the view that scientific hypotheses within naturalistic theorizing about concepts are untenable without the support of reliable and adequate empirical evidence. However, what is at stake is whether metaphysical issues (or, for that matter, non-empirical arguments provided by philosophers) are an irrelevant contribution to assessing the plausibility of the theoretical models posited by psychological theories of concepts.

Having established the distinction between two broad doctrines to the problem of the nature of concepts, we can move on to address the answer that proponents of EpiC and MetC can offer to the ontological question. My position in this respect is not intended to be neutral. As I will argue, theories committed to EpiC fail to answer the ontological question in a way that is pertinent to the representational framework that underpins the project of a scientifically respectable intentional psychology.

2.1 EpiC and the ontological question

Historically, concepts (formerly referred to as ‘ideas’) have been predominantly identified with definitions. Current theories of concepts in cognitive science have been developed in reaction to this traditional view. There are four leading families of theories typically distinguished from one another by the type of structures they postulate lexical concepts to have, given the standard understanding that concepts can be associated with the meaning of individual words in natural languages. These four groups of theories are generically known as the Prototype theory, the Exemplar theory, the Theory-Theory, and Informational Atomism, respectively.

Except for Informational Atomism, all the other theories take it for granted that concepts are complex entities, that is, that they are structured mental representations. In this respect, the models of concepts that these latter theories put forward are not different from the definitional model against which they react. As Margolis (1998) points out, the assumption that concepts have complex structures is ubiquitous among most philosophers and cognitive psychologists interested in the studied concepts:

For example, while the classical theory of concepts says a concept C must decompose into a set of concepts that express necessary and sufficient conditions for the application of C, the prototype theory says that C must decompose into a set of concepts that express statistical conditions that govern the application of C. Similarly, the theory-theory—which is gaining attention in psychological circles—says that a concept C must participate in an inferential system of a certain sort and that C is inherently connected to the other concepts that constitute the system. (p. 347)

There is a relatively wide consensus among cognitive scientists that concepts must have complex structures if they are to account for the explanatory roles many cognitive psychologists have attributed to what they call prototypes, exemplars, and theories (Edwards, 2010b). Notice that the explanatory role of these constructs is
sometimes useful to demonstrate that the class of concepts is a natural kind (Taylor and Vosgerau, 2021), and sometimes it is useful to demonstrate the opposite (Machery, 2009). In any case, the notion of natural kinds in cognitive psychology is preferentially invoked on pragmatic grounds. Moreover, the claim that the main constructs posited by prototype, exemplar, and theory theorists form natural kinds seems to be uncontroversial, even if none of those kinds turns out to be a good paradigm of concepts.

It is far from obvious that the empirical evidence supporting the role of prototype-based, exemplar-based, and theory-based bodies of knowledge in higher cognitive processes can be taken to support the role of concepts in those processes. It is better to say that advocates of PET-BoKs have empirically grounded reasons to claim that their preferred BoK plays an important role in the explanation of such and such epistemic capacities (e.g., rapid or reflective categorization). The reason for this is that there has to be more to the identification of concepts with one of these BoKs than empirical adequacy concerning some salient explanatory functions. Insofar as psychological explanations are intended to generalize over representational states composed of concepts, theorists would be justified in inferring that a given class of BoKs is a candidate for being identified with the class of concepts in case it is likely to satisfy all the functions concepts are expected to serve for the sake of generalization. However, even though there is broad consensus about the challenges that an adequate theory of concepts must meet (Laurence & Margolis, 1999), defenders of PET-BoKs tend to reduce the explanatory roles that their preferred model of concepts is expected to serve (see 3.1.1, below). Once again, the criterion for this reduction is the pragmatic value, or usefulness, of the considered theoretical construct, given the working hypothesis that concepts have the structural characteristics associated with the type of BoKs that figures prominently in the explanation of experimental tasks where some notion of complex mental representation seems to be implicated (e.g., categorization by similarity to a prototype).

Some might think that reducing the set of explanatory functions that a given class of BoKs should serve is a legitimate move if it is done on pragmatic grounds. For example, Contreras (2016) puts forward an argument against (the scientific status of) concepts that hangs upon the distinction between what the author calls theoretical and explanatory desiderata. Admittedly, the strategy does provide a certain latitude for testing new working hypotheses about cognitive phenomena in which concepts are supposed to be involved. However, it is also worth noting that the case for PET-BoKs based on reasons of practicality (notably, researchability) is constrained by a widespread preliminary assumption about the nature of concepts that conflicts with a CRTM and the related project of a science of the mind that vindicates the generality of intentional explanations. According to this assumption, the identity conditions of concepts can be equated with their possession conditions. If this assumption were right, the answer to the question of what a concept is would be reduced to an answer to the question of what it is to have a concept. Unfortunately, it is, at least, unclear that the properties in virtue of which people are attributed the possession of these BoKs can tell us something serious about what concepts are. To the contrary, there are reasons to think that such and underlying assumption is mistaken and even incompatible with the default assumption in cognitive science that concepts are mental particulars (Fodor, 2004)—Nonetheless, see Weiskopf & Bechtel (2004) for some dissenting remarks.

Among cognitive psychologists interested in concepts, it is generally assumed that what a concept is depends on what a concept does (Edwards, 2014). The reason for this is that the possession conditions of concepts are taken to be equivalent to being able to do something or, as Fodor (1998) used to put it, to a kind of

---

6 This contradiction should not be surprising if we consider that the very notion of natural kinds is controversial enough (e.g., Hacking, 2007).
Primitive Concepts and the Ontological Question

Bernardo Antonio Pino Rojas

epistemic ‘know-how’. In fact, theories positing the existence of PET-BoKs focus on what agents like us can do with those BoKs, which typically include our epistemic capacities for categorization, recognition, inference, and the like (Edwards, 2009). Given the pragmatic value that the putative structures of PET-BoKs have regarding those cognitive capacities, researchers have felt tempted to conclude that concepts must be structured entities. And there seems to be good reasons for adopting such a strategy. If theories of PET-BoKs manage to individually show that cognitive processes where concepts are presumably implicated require the manipulation of representations with some particular type of structure, then the problem of the nature of concepts reduces to choosing among the relevant alternatives on offer (notably, among the posits of prototype, exemplar or theory theories). However, there are, at least, two unwanted consequences of this strategy. The first is that, even while it is admitted that complex concepts are made up of other constituent concepts, the nature of the most basic or primitive structural constituents is left unattended. The second is the conclusion that concepts are complex mental representations from the controversial premise that the pragmatic value of appealing to our epistemic capacities is ontologically important. Thus, a serious answer to the ontological question is bracketed or ignored on practical grounds, since we are left with theoretical models of concepts that are gratuitously assumed to be metaphysically dependent on our epistemic capacities. Hence, conceiving concepts as PET-BoKs amounts to an epistemically motivated notion of concepts, in that the satisfaction of certain methodological demands is taken to exhaust that which is ontologically important to the kind of things concepts are.

2.1.1. An explanationist definition of concepts as a case in point regarding EpiC

Among theorists advocating prototype-based, exemplar-based, and theory-based BoKs on explanatory grounds, there is no agree-

ment as to which of those complex mental representations can be said to designate the class of concepts (Machery, 2010). In light of this disagreement, some authors have considered the reasons for an eventual splitting of the class of concepts, adopting the notion of natural kinds as homeostatic property clusters developed by Boyd (1989, 1999). For instance, Piccinini & Scott (2006) have claimed that a proposal for concept splitting must meet the following requirements:

[… it must be shown that the phenomena antecedently associated with concepts divide into different, largely independent groups; that these groups of phenomena are explained by different kinds of internal structures possessing different clusters of scientifically relevant properties; and possibly that the cooccurrence of those clusters is explained by different causal mechanisms. (p. 393)

What matters for present purposes is that to do the stated task properly, these authors think that an adequate programmatic definition of the notion of a concept is needed. In their view, if we want to seriously determine whether concepts split into different kinds, we need to count on a more restrictive notion of a concept. My interest in this restricted definition is that it presents a clear endorsement of the claim that issues about the nature of concepts derive from the pragmatic value of PET-BoKs as mentioned above. According to Piccinini & Scott (2006), for a given type of mental entity to count as concepts, they must be identified with homeostatic property clusters that are centrally involved in the explanation of some set of cognitive phenomena. In their words,

Within psychology, concepts are mental particulars posited as a specific part of the explanation for certain cognitive phenomena. For present purposes, a concept is anything that plays the relevant role within the explanation of those phenomena. Thus understood, concepts are a singular natural kind if and only if all such entities turn out to share a large set of scientifically relevant properties as a result of the same causal mechanism. (pp. 395-396)
Then the authors proceed to specify what they call a programmatic definition of concepts, according to which concepts are “the constituents of structures posited in the explanation of certain psychological phenomena” (p. 396). To adapt a familiar analogy, if pumpkins are theoretical structures whose processing accounts for, say, categorization, then concepts are the constituents of pumpkins. For a programmatic definition, nothing else is regarded to be relevant to say about those constituents. The list of phenomena in question involves discrimination, nonlinguistic inference, categorization, word and sentence understanding, linguistic inference, and lexical combination. Piccinini & Scott (2006) suggest that an assessment of whether the class of concepts form a natural kind should be carried out concerning this set of psychological phenomena.

Key to this strategy is the idea that an argument for concept splitting must show that different kinds of mental representations are needed to explain different phenomena where concepts were previously thought to play explanatory roles. Hence, to the extent that PET-BoKs are models of concepts that fail to independently account for the proposed desiderata on concepts, the conclusion that the class of concepts split into different natural kinds appears to be inevitable. Authors such as Machery (2009) have explicitly argued that debates over the nature of concepts as a single natural kind among prototype, exemplar, and theory theorists are beside the point because these theorists’ preferred constructs have few properties in common and may be independently required in the explanation of the different relevant phenomena.

Concept theorists convinced by the argument for the heterogeneity of PET-BoKs might feel inclined to concede that concepts split into different natural kinds because these BoKs do. However, this conclusion should only seem appealing to those who agree to exclude from the list of desiderata two other phenomena that Piccinini & Scott (2006) consider irrelevant to a purely psychological theory of concepts (i.e., a theory of some BoK), namely reference and publicity. Roughly, reference has to do with the relation obtained between a concept and the things in the world to which the concept applies correctly, and publicity has to do with the property that concepts presumably have if they are to be shared by different people or the same person at different times. To someone committed to a CRTM, on the other hand, it stands to reason that a complete theory of concepts has something to say about the publicity of concepts and the stability of conceptual content, since both properties are thought to be required for scientific generalizations that quantify over mental representations (type). It would be arbitrary simply to take them off the table because they do not conform to some practical criteria of researchability once it has been established that PET-BoKs are too heterogeneous to form a single natural kind. Thus, even if we granted that PET-BoKs are distinct psychological entities, the conclusion that the very class of concepts splits into distinct kinds is still in need of further justification.

In what follows, I will focus on a philosophical perspective that takes issue with the widely shared assumption that concepts are metaphysically dependent on their relations to other concepts. I aim to show that, although atomistic theories of concepts are both unpopular and embryonic within cognitive science, an atomistic approach to concept individuation is in a better position to avoid the most serious problems inherent in the view that concepts are complex mental representations. If so, it is debated about whether PET-BoKs are distinct natural kinds that are beside the point when trying to answer the question of what concepts are.

---

*I take the type/token distinction regarding mental representations for granted. However, I will avoid invoking the distinction unless it is necessary for clarity of exposition.*
2.2. MetC and the ontological question

The central claim in this section is that the development of a theory of concepts with a higher chance of being a successful theory (i.e., an approximately true scientific theory) than its rivals cannot be carried out solely based on its empirical adequacy. The reliable procedures of experimental research are insufficient to discriminate which of the current psychological theories of concepts offers a plausible story about the constitutive properties of concepts, especially in the context of ill-advised theoretical proposals.\(^8\)

Admittedly, philosophers’ intuitions and non-experimental reflections on the metaphysics of concepts are not regularly regarded as relevant for evaluating empirical theories in cognitive psychology. However, experimentally based theories of concepts inevitably contain implicit assumptions about the nature of concepts. To the extent that these assumptions are either taken for granted or unaddressed by cognitive scientists, philosophers can contribute to the choice between rival theories by making those preliminary assumptions explicit and assessing their metaphysical plausibility. This is more or less the role of philosophy regarding the study of concepts defended by Vallejos (2008) that I intend to consider for characterizing a metaphysically motivated approach to the ontological question.

As argued elsewhere (Fodor, 1998, 2008; Higginbotham, 1998; Rey 1985, 2010), concepts must be distinguished from the epistemic conceptions one might have about them or the things that fall in their extension. While PET-BoKs are said to figure prominently in psychological generalizations regarding several cognitive competencies, each of these BoKs has been seriously questioned as an adequate central posit of a theory of concepts for its reasons. Common to all of them is the failure to cohere with the fact that concepts transcend the idiosyncratic variability of our epistemic conceptions (Rey, 2010). Nonetheless, a good theory of concepts should ultimately allow for both an account of the position of prominence that PET-BoKs have in psychological tasks and a notion of concepts that is immune to the insurmountable problems of its main contenders. That is, this notion should avoid the problems that stem from the view that concepts must be complex mental representations in virtue of adopting EpiC.

To begin with, consider that the empirical success of theories that identify concepts with complex mental representations depends on pragmatic criteria such as their explanatory power. Theories in the same field are expected to explain certain phenomena, and the one that explains them more effectively is said to have more explanatory power. However, even though the progress of scientific inquiry in special sciences\(^9\) is linked to the evaluation of theories in terms of criteria related to explanatory adequacy (including, e.g., simplicity, generality, coherence with other theories, explanatory power, etc.), philosophers’ non-empirical theory assessment is required to examine how certain preliminary assumptions about the ontology of concepts encourage methodological constraints that the theories are compelled to satisfy.

A case in point is when theories committed to highly conflicting assumptions (e.g., a similarity-based theory of concepts versus a theory-based theory) afford considerably different degrees of latitude in terms of the properties that they take to be constitutive of concepts (e.g., representation of certain superficial features of category members versus representation of bodies of causal knowledge underlying superficial regularities). Since an adequate theoretical notion of a concept is expected to serve various func-

\(^8\) Given the preliminary working assumption regarding CRTM stated in the introduction, I take it for granted that an adequate theory of concepts must be consistent with the traditional goal of unraveling the mental representations required for counterfactual-supporting generalizations in the scientific explanation of observable behavior (Fodor, 1998; Cain, 2002).

\(^9\) The expression ‘special sciences’ is normally used to talk about sciences other than physics, including psychology.
tions, cognitive scientists’ best provisional explanations appear to be a criterion for the reality of the entities designated by models of concepts with certain structural characteristics (Taylor and Vosgerau, 2021). However, if empirical adequacy were the only standard for assessing the ontological plausibility of a theoretical posit, highly heterogeneous models with similar explanatory targets and effectiveness would be a reason for questioning the coherence and the unifying role of the construct (e.g., Machery, 2010). It is, therefore, all but uncontroversial to claim that, because the explanatory desiderata are established on an \textit{ad hoc} basis by theories committed to conflicting assumptions about the nature of concepts, the empirical effectiveness of the theories is (most of the time) incapable of being considered comparatively.\footnote{This situation is even more evident when one considers different views about the nature of concepts that fall outside the scope of the current paper. See, e.g., Peacocke (1992), Dummet (1993), and Prinz & Clark (2004).} The non-experimental task of exposing and analyzing these kinds of assumptions is by no means a futile philosophical contribution, especially when they regulate the empirical scope of a theory and the significance of evidence that is deemed to support it (Aguilera & Pino, 2019). Hence, the suggestion that the outcome of an ontological analysis of concepts must count as an additional argument to the best explanation.

Following Vallejos (2008), the philosophical undertaking of elucidating the nature of concepts requires formulating conditions the satisfaction of which allows for the individuation of properties that are constitutive of concepts. This task involves establishing at least three related issues, namely, the identity conditions of concepts, i.e., what kind of thing a concept is; the individuation conditions for the semantic properties of concepts, i.e., in virtue of what kind of properties a concept has meaning; and the possession conditions of concepts, i.e., what it is to have a concept (Rodríguez, 2007). Here the order of explanation matters, as is evident in polemics about the metaphysical dependence of concepts on certain epistemic capacities that are presumably constitutive of the conditions for possessing concepts (Fodor, 1998, 2004; Weiskopf & Bechtel, 2004). Drawing on these discussions, I take the problem of the identity conditions of concepts to be before any other question when it comes to establishing the kind of things that concepts are. If this is correct, appealing to the pragmatic value of our epistemic capacities is irrelevant to the ontological question. Consequently, it is a constraint on any acceptable MetC that the three types of conditions be coherently approached from an ontological perspective (i.e., nonepistemically).

As argued in the case of EpiC, the conclusion that concepts can be identified with complex mental representations hangs on the exclusion of reference and publicity from the explanatory goals of a psychological theory of concepts. However this exclusion is arbitrary and can be better understood in terms of BoK-based theories’ lack of resources to account for the intrinsic representational properties of concepts (Pino & Aguilera, 2017). Consider that these theories are committed to \textit{Inferentialism} regarding concept possession, according to which possessing a concept involves being able to draw the right sorts of inferences (Weiskopf 2013). However, since those inferential relations (or, for that matter, bodies of knowledge) are subject to inevitable variation either between people or over different stages of a person’s life, the possibility of establishing robust identity conditions for concepts is hindered. Rey (2010) has raised a similar concern over BoKs in terms of what he calls the problem of \textit{epistemic variability}:

\begin{quote}
Unless one restricts the relevant knowledge in some principled way, no two people (or stages) will share a concept, since, short of coincidence, no two people (or stages) will bring exactly the same knowledge or procedures to bear in making many of their judgments. (p.221)
\end{quote}
Thus, insofar as theories of concepts positing PET-BoKs do not have the resources to explain how the content of concepts can remain stable across changes in beliefs, the problems of epistemic variability and unstable identity conditions can be said to be inherently linked to that kind of theories (Fodor & Lepore 1992). By contrast, in the philosophical perspective addressed in this section, the suggestion is that issues about concept identity should be taken to be metaphysically independent of thinkers’ epistemic capacities, irrespective of the crucial role concepts presumably play in the explanation of our cognitive capacities.

The question arises as to whether non-experimental philosophy has a role to play in current theorizing about concepts within the framework of CRTM. I claim that it does, and, by way of example, the following section turns to conceptual atomism with a focus on Fodor’s account of the basic constituents of thoughts, which is a key component in his broader endeavor to develop scientifically respectable intentional psychology.

2.2.1. Informational Atomism as a case in point regarding MetC

Informational Atomism (IA) is a theory of concepts developed by Fodor (1998) in reaction to inferentialist theories that identify concepts with structured mental representations. According to IA, typical concepts are unstructured mental representations whose (type-)identity conditions do not depend on how they are related to other aspects of cognition, but on how they are reliably locked to things in the world. I previously noted that a tenable account of concept identity was an unsurmountable obstacle for theories advocating models of structured concepts. By contrast, since IA construes conceptual content nonepistemically, it is immune to the problems of content and epistemic variability as faced by inferentialist contenders. Fodor’s theory is not only worth our attention for providing a plausible story about how concepts can be shared but also for being a version of MetC capable of satisfying explanatory demands.

IA is committed to a causal theory of content according to which there is a nomic relation between instances of symbol M and the property $m$ that M expresses. Because of this relation, the tokens of the symbol are said to carry reliable information about the property that causes them. Margolis (1998) depicts the point in the following way:

Put in mental terms what this means is that the concept BREAD expresses the property bread because bread is the reliable cause of BREAD-tokenings. In other words, there is a law connecting the property of being bread with the property of being a BREAD-tokening, and it is in virtue of this law that the concept BREAD expresses the property bread. (p. 350)

An atomistic theory of concepts, combined with an informational theory of content, makes it possible to establish an immediate contrast between the type of information that is essential to the concept and the type of information that is not. For example, in the case of the concept BREAD (i.e., a mental representation type), the essential information is that to which, by natural law, the tokens of that concept (i.e., mental particulars with semantic and causal properties) are locked. Any other information related to the concept BREAD is not essential to it, such as the piece of information that too much white bread can contribute to obesity. Here we have a clear contrast with theories that postulate the existence of PET-BoKs, where that which determines the content of a concept is construed either in terms of many inferential relations (e.g., the theory-theory) or in terms of less than many (e.g., the prototype theory). The distinction between essential and non-essential information becomes at least problematic in the case of those theories, unlike the case of conceptual atoms causally controlled by their referents.

An account of a concept’s identity in virtue of mind-world relations, together with the realization that theories identifying concepts with complex mental representations in cognitive science are likely to be untenable, constitute the strongest advantage of an
atomistic alternative. Despite this advantage, conceptual atomism has met severe criticisms which have led many to dismiss it as a non-starter, even when recurrent criticisms can be countered by showing that they are often made on irrelevant grounds (Edwards, 2009; Kwong, 2006). Let us consider four recurrent objections:

**The problem of error.** Theories of concepts are expected to account for people’s incorrect application of the concepts they possess, such as, for example when a person mistakes a faux fur cushion for her pet cat— as it might be. In such situation, an information-based theory of content would appear to be unable to account for the incorrect application of the concept CAT, given that it is cats, not cushions, that are supposed to be the reliable causes of CAT-tokenings. A plausible solution that Fodor (1990) gave to this problem is known as the asymmetric-dependence theory. According to this theory, while there is a lawful relation between a concept (e.g., the concept CAT) and the property that it expresses (e.g., the property of being a cat), relations between that same concept and other properties (e.g., the property of being a cushion) would be asymmetrically dependent on the former relation. In other words, the idea is that the incorrect application of a concept would not hold were it not for the lawful relation that governs its correct application, but not vice versa—Hence, the asymmetric dependence. Thus, IA has the resources to explain, e.g., how the meaning of the atomic concept CAT can remain the same, even when it is caused by cushions that look like cats under certain circumstances.

**Explanatory impotence.** Theories that identify concepts with complex mental representations rely on the putative structural components of concepts to explain such psychological phenomena as categorization, inference, and so forth (Prinz, 2002). In contrast, since conceptual atomism identifies concepts with unstructured mental representations, it seems to lack the resources to do an important explanatory work. By and large, criticisms about the explanatory impotence of conceptual atomism are linked to the identity between the individuation of conceptual content and the individuation of primitive concepts that IA appears to posit (Rodríguez, 2007). Once we set aside such incorrect interpretations, the criticisms can be countered in different ways. For example, Kwong (2007) argues that, by dissociating conceptual atomism from what he conceives as radical or untenable views about content determination and the possibility of punctate minds, it is possible to defend a theory that resists charges of explanatory impotence. Here the main idea is that phenomena such as different types of categorization judgments could be explained by appealing to information that is not contained in concepts but collateral to them. Collateral information (e.g., the information PET-BoKs are said to encode) can thus be understood in the form of collateral structures associated with atomic concepts, on the understanding that all information about a concept is always collateral—cf., Laurence & Margolis (1999).

**The problem of coreferential and empty concepts.** It has been objected that IA is unable to make sense of certain types of concepts, namely those that fail to refer to, such as UNICORN or SANTA CLAUSE. This problem has been considered inherent in any atomistic theory of concepts that excludes inferential relations as constitutive elements of the content of concepts. For critics of IA, unless inferential relations involved in cognitive processes are also treated as constitutive of the content of concepts, it is hardly credible that a theory appealing to nomic relations holding between concepts and properties in the word has the resources to distinguish between coreferential concepts in general, and concepts that fail to refer, in particular.

An overworked example involves the properties of being water and being H2O. Given that these properties are generally thought to be identical, the problem arises as to how a reference-based
theory of concepts can distinguish between the concepts WATER and H2O, given that they are nomically controlled by the same property. Likewise, since so-called empty concepts such as NESSIE and KRYPTON have the same reference (i.e., they refer to nothing), IA seems to inevitably lead to the conclusion that all empty concepts have the same content. But this conclusion is strongly counterintuitive since many empty concepts (e.g., HAMLET, PHLOGISTON, and MACONDO) are distinct concepts. Again, an advocate of atomism can respond to the objection by showing that IA may have the resources to provide a plausible account. Indeed, Edwards (2010a) has proposed a reference-based approach to concepts that capitalizes on “the epistemic and practical situation of the concepts’ possessors” (p. 115) to suggest a tentative account of empty concepts. Thus, while many believe that the challenges posed to IA by coreferential and empty concepts show that a working atomistic theory must allow for some concepts to be distinguished by virtue of certain inferential role semantics (Laurence & Margolis, 1999), there is still more to say about the relationship between a concept’s content and the explanatory role of certain non-concept-constitutive factors.

Conceptual nativism. CRTM relies on the existence of a finite set of primitive concepts out of which infinitely many complex concepts are constructed, and it is said to be a problem for IA that concepts like FLUTE, BALLOON, and CAR be part of that primitive stock if they are required to be innate. However, nativism regarding concepts is encouraged by the realization that there is no plausible explanation for concept learning based on hypothesis confirmation (Piatelli-Palmarini, 2018). If learning a concept implies representing a hypothesis where the concept to be learned is already involved, then it is incoherent to talk about concept learning (Piatelli-Palmarini, 2018; Weiskopf, 2013). If this is correct, and if learning as hypothesis confirmation is the only plausible model of learning that is consistent with CRTM, IA is in want of a credible story about the acquisition of primitive concepts. Nevertheless, while it is a fact generally overlooked that the task of specifying which concepts are innate is an empirical one (Laurence & Margolis, 1999), advocates of IA can avoid radical conceptual nativism by appealing to either non-psychological processes or psychological processes that are not constitutive of the identity conditions of concepts. In the first case, for example, Fodor (1998) suggests a solution according to which our innate sensorium enables the concept of an artifact (e.g., FLUTE) to become locked onto its referent (i.e., flutes) probably by some natural law. Crucial to this non-cognitive solution is the contention that it is wrong to conclude that primitive concepts are innate from the realization that no concepts can be learned (Fodor, 2008, chapter 5). The reason is that there is a myriad of possible ways in which one could acquire concepts from experience without learning them using hypothesis formation and confirmation (e.g., when the use of some experimental equipment mediates the causal interaction between the tokens of a concept and tokens of its referent). An example of the second case is the proposal advanced by Margolis (1988) in terms of sustaining mechanisms (e.g., bodies of information about salient properties of squirrels) which are not constitutive of the identity conditions of concepts but provide reliable indications of the properties that can cause tokens of a concept (e.g., SQUIRREL). However, as discussed below, an adequate solution to the problem of conceptual nativism must be compatible with the formulation of conditions for possessing concepts in non-epistemic terms (Vallejos, 2008).

Let’s now turn to the issue of whether MetC can shed interesting metaphysical light on the problem of the nature of concepts. It should be admitted that IA is notoriously in need of further theoretical development, partly due to the relatively little attention that atomistic theories have received among philosophers and cognitive scientists. Because IA reacts against current views of conceptual structures without yet offering a clear criterion for
individuating unstructured concepts, it has been considered a negative view (Laurence & Margolis, 1999). However, while theories that identify concepts with complex mental representations face problems that make them unworkable, it remains to be seen whether the development of an atomistic mode of individuation of concepts meets a similar fate (Vallejos, 2008). If this is not a good enough reason to subscribe to an atomistic view of concepts, it is at least a good reason to argue that IA is to be preferred as the last standing theory because it promises to provide a coherent answer to the ontological question. As previously stated, the problem of the nature of concepts involves formulating three types of conditions, and MetC requires that each of them be approached nonepistemically:

(1) Conditions the satisfaction of which allows for concept individuation

(2) Conditions the satisfaction of which allows for the individuation of the semantic properties of concepts

(3) Conditions the satisfaction of which allows us to determine what it is for someone to possess a concept

Insofar as IA identifies primitive concepts with atomic mental representation types capable of being instantiated in people’s minds, it allows for concepts’ identity conditions to be established nonepistemically. That is, it allows for concept individuation to be specified independently of any other aspects of cognition (e.g., psychological processes, inferential relations between concepts, perception, etc.). In turn, the informational component of the theory allows for the content of concepts to be construed essentially. This account of the semantic properties of concepts is consistent with the requirement that people’s beliefs need not be constitutive of the concept. Indeed, IA states that possessing a concept is not constituted by inferential relations or some type of BoKs, but by a person’s possessing the relevant mental particular. A positive effect of this account of concept possession is that it allows for the distinction between concepts and the beliefs (or conceptions) that we may have of them. Confusing the two has proved to seriously hinder the development of a working theory of concepts (Higginbotham, 1998; Kwong, 2007).

3. Concluding remarks: What could PET-BoKs be if not concepts?

I have argued that an atomistic mode of concept individuation may be in a better position to answer the ontological question. This contention is based on the type of theory evaluation that I have endorsed, using which it is possible to expose the difficulties encountered by a theory of concepts that seeks to establish itself as the best among its rivals merely on pragmatic criteria of comparison. To illustrate the contribution of philosophy to the scientific study of concepts, I have distinguished between two broad approaches to the ontological question, one that appeals to an epistemically motivated model of concepts (EpiC) and the other that appeals to a metaphysically motivated model of concepts (MetC). In consonance with the polemical position regarding concepts defended by Fodor (1998), the distinction allowed me to show that an argument for the conclusion that concepts are complex (or structured) mental representations from the premise that the pragmatic value of appealing to our epistemic capacities is ontologically important is invalid.

Someone could legitimately question the strategy of favoring an answer to the ontological question about concepts to defend an atomistic option. After all, the referents of the epistemically motivated theoretical constructs in some of the rival theories seem to correspond to a series of real cognitive mechanisms. If

---

11 I would like to thank another anonymous referee for pointing me to this possible concern.
so—this person might think,—perhaps, it is preferable to seek other epistemically motivated options, even in the current state of disarray of the field. However, for a scientifically respectable theory, it is not enough to merely establish whether any of the competing notions of concepts refer to a possibly real entity, but also to establish whether such a notion is in a better position to pick out the type of real entity a plausible theory of concepts is expected to be about (which is ultimately a matter of discovery). And, as initially stated, a theory of concepts is in a better position to so when it is consistent with a background theory like CRTM, even if this background theory is a remotely plausible theory of the cognitive mind. It has been part of the polemical strategy of the present paper that an answer to the ontological question about concepts is particularly relevant in the current state of confusion of the study of concepts. Indeed, by distinguishing between EpiC and MetC, we have not merely illustrated a tension between two different approaches to the problem of the nature of concepts. Rather, we have established that a comprehensive solution to this problem is unviable when solely relying on explanatory factors inherently based on contingent epistemic mechanisms.

If, as a result of the philosophical contribution outlined in the previous section, one feels inclined to embrace some version of conceptual atomism, then one may also feel drawn to rejecting PET-BoKs as adequate models of primitive concepts. Of course, the reader should resist the temptation to conclude that the atomistic alternative is a better theory solely because of its non-epistemic character. The point is that, while the epistemic theories that dominate the current theorizing about concepts face problems that make them untenable, atomism is a coherent option that awaits further theoretical development. Still, an important body of evidence in empirical psychology is taken to support the existence of those BoKs. Thus, the question arises as to what PET-BoKs could be if not concepts.

There are certain constraints on an adequate theory of concepts that follow from a commitment to CRTM (see, Fodor, 1998, Chapter 2, for further details). A fundamental constraint is that concepts must be mental particulars. I hope I have made the case that PET-BoKs fail to meet this condition due to their intrinsic cognitive character. Nonetheless, in the psychological literature on concepts, an important body of empirical evidence supports the relevance of BoKs conveying some type of characteristic information in our mental life. Given that, our question about what these BoKs are if they are not concepts reduces to the question of how to reinterpret the evidence in favor of these BoKs and the place they might occupy in a theory of cognition. Presently, we lack a detailed alternative, but there are a few sketchy options that illustrate the viability of the task at hand and suggest a natural progression of the present work. One option, for instance, construes PET-BoKs in terms of sustaining mechanisms involved in the acquisition of concepts. In this case, PET-BoKs can be understood as structural entities that are somehow associated with but not constitutive of concept identity (Kwong, 2007; Margolis, 1998). Another option defends the view that reference is the fundamental function of a concept and suggests that the functions often attributed to constructs such as PET-BoKs can be plausibly construed in terms of the implementation of such reference relation (Edwards, 2011, 2014). Overall, the moral is that polemics in favor of an important reinterpretation of PET-BoKs, insofar as they are motivated by a viable version of conceptual atomism, add plausibility to an ontologically inclined philosophical contribution to the scientific study of concepts.
REFERENCES


Primitive Concepts and the Ontological Question

Bernardo Antonio Pino Rojas


Primitive Concepts and the Ontological Question

Bernardo Antonio Pino Rojas


