Overview

Philosophical issues about concepts

Joseph McCaffrey and Edouard Machery*

In this article, we review some important controversies about concepts in the philosophy of psychology, focusing particularly on the theories of concepts developed in philosophy, on the debate about the homogeneity of concepts, on neo-empiricism, and on concept learning. © 2012 John Wiley & Sons, Ltd.

How to cite this article:
WIREs Cogn Sci 2012, 3:265–279. doi: 10.1002/wcs.1166

INTRODUCTION

In this article, we review some important controversies about concepts in the philosophy of psychology. We focus first on what concepts are, before examining the theories of concepts developed in philosophy. In the section on *Are Concepts a Natural Kind?* we examine the recent debate about whether concepts form a natural kind and about whether the theoretical term 'concept' should be eliminated from the theoretical vocabulary of psychology. In the section on *Neo-Empiricism*, we examine the empirical and theoretical arguments for neo-empiricism. Finally, we examine the philosophical and psychological debates about concept learning.[a]

WHAT ARE CONCEPTS?

Theories of concepts have been developed in philosophy[b] and in cognitive science[c] (including in cognitive neuroscience, AI, psychology, etc.), and it is commonly assumed, particularly by philosophers such as Georges Rey, Eric Margolis, Steve Laurence, and Jerry Fodor, that these theories aim at resolving the same kind of issues. In addition, many philosophers hold that the theories of concepts developed in cognitive science fail to resolve the issues of interest among philosophers. For instance, Fodor concludes his review of Gregory Murphy's book on the psychology of concepts, *The Big Book of Concepts*, as follows (Ref 1, p. 4):

> It is part of our not knowing how the mind works that we don't know what concepts are or what it is to have one. Just about everything that current cognitive science says about either topic is wrong. But at least it is clear that concepts aren't typicality structures and that having them is not being able to sort things. Except for leaving that out, Gregory Murphy's book tells you most of what there is to the psychology of concepts. Read it, therefore, by all means; but don't even consider believing it.

On the other hand, some philosophical theories of concepts have been criticized for failing to explain what cognitive scientists seem to be concerned with, such as how we classify objects into kinds, how we draw inferences, and so on (see, e.g., Prinz's criticism of Fodor's atomistic theory of concepts in Ref 2).

Philosophers have failed to convince cognitive scientists of their errors, and vice versa. This failure suggests that, when they criticize each other, philosophers and cognitive scientists are speaking at cross-purposes and that the theories developed in philosophy and in cognitive science are not meant to resolve the same issues and thus are not competing.

But, then, what issues are theories of concepts in philosophy and cognitive science concerned with? The most influential theories of concepts in philosophy such as Fodor's or Peacocke's have been concerned with explaining how we are able to have propositional attitudes (beliefs, desires, etc.) about the objects of our attitudes. Philosophers interested in concepts typically want to determine the conditions for having beliefs about, say, dogs as such or desires about, say, burgers as such. And when a philosopher like Fodor warns against the theories of concepts developed in cognitive science, it is because they fail to explain how we can have propositional attitudes about the objects of our attitudes. It is however patent that the theories of concepts developed in cognitive science typically

*Correspondence to: machery@pitt.edu
History and Philosophy of Science, University of Pittsburgh, Pittsburgh, PA, USA

have different goals. Cognitive scientists want to explain why people categorize, draw inductions, make analogies, combine concepts, etc., the way they do. For instance, they want to explain why inductive judgments are often sensitive to similarity (e.g., Ref 3; for review of the psychology of induction, see Ref 4, chapter 8; Refs 5 and 6). Thus, the theories developed by cognitive scientists and philosophers rarely attempt to resolve the same issues.[d]

Taking this conclusion for granted, one could endorse two different views about the relation between the theories of concepts developed in philosophy and in cognitive science. One could propose, as Carey,[7] Lombrozo,[8] and Margolis and Laurence[9] do, that these theories focus on different aspects of the same entities. Instead, one could propose, as Machery[10,11] has done, that 'concept' refers to different entities when used by psychologists and cognitive scientists.

THINKING ABOUT

In virtue of what are we able to think about, e.g., dogs as such? Philosophers tend to favor one of two possible responses to this question. According to inferentialists,[12,13] our capacity to think about dogs as such is in some way connected to the inferences we are disposed to draw. According to some versions of inferentialism, roughly, one is able to think about dogs if (and on some views if and only if) one is disposed to draw some specific inferences: For instance, thinking about dogs as such (instead of thinking about cats or about canines) may involve being disposed to draw some inferences about an organism (e.g., to infer that it is an animal or that it is likely to bark) from the fact that it has been classified as a dog; thinking about dogs as such (instead of thinking about cats or about canines) may also involve being disposed to infer that an organism is a dog from the fact that is has some properties (e.g., that it barks, has a specific shape, etc.). According to other versions of inferentialism (particularly, Peacocke's[13] theory), one is able to think about dogs if and only if one is one is disposed to find some particular inferences to be 'primitively justified', that is, roughly, to be justified merely in virtue of the fact that these inferences contain the concept of interest. This idea is best illustrated by logical concepts such as AND.[e] According to Peacocke (Ref 13, p. 6), for instance, to be able to entertain conjunctive thoughts (viz. to have the concept AND) is to be disposed to take the following types of inferences to be justified merely in virtue of the fact that they contain AND:

1. p and q
   p
2. p
   q
   p and q

Inferentialist theories disagree about how to identify the inferences that are constitutive of particular concepts. Holists such as Ned Block[14] hold that all the inferences a concept features in are constitutive of this concept, while molecularists such as Peacocke hold that only some particular inferences are constitutive of a concept. Of course, molecularists take on the burden of specifying, in a principled manner, which inferences are constitutive of particular concepts. To illustrate molecularism with the concept SQUARE, Peacocke (Ref 13, p. 74–90) has argued that to possess this concept it is necessary to consider an inference from the perception of a square object presented in a particular, canonical orientation, to the indexical belief that this [perceived object] is square to be primitively justified.

Inferentialist theories of concepts stand in sharp contrast with atomism about concepts (mostly associated with Fodor, see Refs 15, 16, and 17). According to atomism, our capacity to think about dogs as such—viz. the possession of a concept of dog—is in no way connected to the inferences we are disposed to draw. Rather, according to Fodor, we are able to think about dogs as such (we have a concept of dogs) because some particular mental representation stands in a nomological relation with the property of dogness: Ceteris paribus, whenever a dog is perceived, we entertain this mental representation. On this view, it is possible to possess a single concept (e.g., the concept of dog)—whence the name 'atomism'—while for inferentialists the possession of any particular concept supposes that one possesses a set of related concepts too (because inferences always involve several concepts).

How are we going to decide between these views? Philosophers often hold that the different theories about concepts have to be consistent with a number of alleged facts about concepts and thoughts[2,16]:

1. *Publicity*: Several individuals and the same individual at different points of time can have the same concept.
2. *Productivity*: Any individual can entertain an infinite number of concepts and thoughts.
3. *Systematicity*: The capacity to entertain some thoughts (e.g., that Mary loves John) comes with the capacity to entertain some structurally related thoughts (e.g., that John loves Mary).

Systematicity is meant to show that concepts combine: Because the thought that Mary loves John is made of the concepts of John, of Mary, and of love according to a particular way of combining concepts, whoever can form this thought by combining the relevant concepts can also form the thought that John loves Mary since the same concepts and the same way of combining concepts are involved. Productivity is meant to show that concepts combine recursively.[18–20]

In addition, some philosophers hold that theories of concepts have to be consistent with some further alleged facts about concepts and thoughts, which are however more controversial.

1. *No analytic/synthetic distinction*: It is not possible to distinguish between the beliefs that are true (or justified) only in virtue of the concepts they involve and the beliefs whose truth (justification) also depend on facts of the world.
2. *Reverse compositionality*: Whoever possesses a concept that can be formed by combining other concepts (e.g., PET FISH) also possesses these other concepts (e.g., PET and FISH).

Proponents of the analytic/synthetic distinction hold that a belief like the belief that bachelors are unmarried is true (or, under some versions of this distinction, justified) only by virtue of the meaning of BACHELOR, while the truth (or justification) of a belief like the belief that bachelors are untidy also depends on empirical facts (for review see Ref 21).

Some accounts of concepts have been judged inconsistent with some of these alleged facts about concepts and thoughts. To illustrate, holism is often taken to be inconsistent with the publicity of concepts: Because any two individuals (or the same individual at two different times) are likely to be disposed to draw somewhat different inferences involving a particular concept (e.g., DOG), they cannot share the same concept if all the inferences involving a concept are constitutive of this concept.[18] By contrast, molecularism does not suffer from this problem since different individuals can be disposed to draw the specific inferences that are constitutive of a particular concept (or to find them primitively justified).

In addition, Fodor has developed a detailed criticism of all inferentialist theories of concepts (including molecularism) by appealing to publicity, the idea that concepts combine (justified itself by appeal to productivity and systematicity), and a rejection of the analytic/synthetic distinction.[15,16,19] Inferentialist theories can either be holistic or molecularist. Holistic theories are false because they are incompatible with the publicity of concepts, as we have just seen. Molecularist theories of concepts need to distinguish those inferences that are constitutive of a concept from those that are not. While this could be done by appealing to the analytic/synthetic distinction—viz. by suggesting that inferences that are analytically truth-preserving (or justified) are constitutive of particular concepts—there is no analytic/synthetic distinction. All the other ways of drawing the sought-after distinction are incompatible with the idea that concepts combine because they entail that one could have two concepts (e.g., PET and FISH) without being able to entertain the complex concept (e.g., PET FISH; this argument is known as the pet fish argument). Suppose for instance that one holds that to entertain any concept is to be disposed to recognize its instances in favorable circumstances (viz. one holds that concepts are 'recognitional'): For instance, to have the concept PET is to be able to recognize pets as pets in favorable circumstances and to have the concept FISH is to be able to recognize fish as fish in favorable circumstances. The problem for this proposal is that being able to recognize a pet and being able to recognize a fish does not guarantee that one is able to recognize pet fish as such: To be able to distinguish pet fish (e.g., goldfish) from things that are not pet fish, one needs more than the capacities to recognize fish and pets; namely, one needs to know what pet fish look like in the world.

Appealing to these alleged facts (e.g., publicity) about concepts has not been enough to settle the controversy about what is necessary or necessary and sufficient to think about something as such (e.g., about dogs as such). First, the controversial nature of certain facts undermines their dialectical weight. For instance, while following Quine[22] many philosophers reject all the versions of the analytic/synthetic distinction, a substantial number of philosophers and linguists disagree.[23–26] Second, even some of the less controversial alleged facts about concepts and thoughts (A–C) have been challenged. For instance, some have suggested replacing the publicity of concepts with the requirement that concepts be such that it be possible for two individuals or for the same individual at different times to have similar concepts (instead of the same concept).[f] Third, even granting the alleged facts about concepts and thoughts, it is dubious whether these provide enough constraints to really winnow the competing theories of concepts. Consider, e.g., Fodor's argument about inferentialism discussed above. This argument fails because it assumes that inferentialists are committed to the homogeneity of concepts, that is, that they hold that *all* concepts are of the same kind. Thus, in the

example given above, it is assumed that all concepts are recognitional—viz. that to possess any concept involves being able to recognize its instances. If one rejects this homogeneity assumption, then one can hold that to possess, e.g., FISH and PET, but not PET FISH, involves being able to recognize their instances. Then, the fact that one may not be able to recognize pet fish while being able to recognize pets and fish raises no particular issue for an inferentialist who holds that concepts combine (and thus that whoever can think about pets as such and about fish as such can *ipso facto* think about pet fish).

In our opinion, the debate about the necessary (or necessary and sufficient) conditions for being able to have propositional attitudes about the objects of our attitudes has reached a dead-end (Ref 10, chapter 2), and it is currently unclear where the way forward lies.

ARE CONCEPTS A NATURAL KIND?

In the remainder of this article, the term 'concept' is used to refer to the bodies of knowledge that are used by default in the cognitive processes underlying our higher cognitive competences (see Ref 10, chapter 1 for this characterization of the notion of concept), and not to whatever it is that allows us to have propositional attitudes about the objects of our attitudes, as we did in the section[g] on *Thinking About*.

Since the early days of psychology,[27] psychologists interested in higher cognition have attempted to study the properties of these bodies of knowledge. They have typically assumed that these share many properties, and that a satisfying theory of concepts should describe these properties, an assumption well captured by Murphy (Ref 4, p. 2–3):

> The psychology of concepts cannot by itself provide a full explanation of the concepts of all the different domains that psychologists are interested in.... The details of each of these must be discovered by the specific disciplines that study them.... Nonetheless, the general processes of concept learning and representation may well be found in each of these domains.

Since the 1970s, cognitive scientists have developed a range of theories meant to capture the properties common to concepts, including prototype theories, exemplar theories, and theory theories (for review, see Refs 4; 10, chapter 4; and Ref 28). In a nutshell, prototypes are bodies of statistical knowledge about the typical or diagnostic properties of a category, a substance, a type of event, etc.[29,30] Prototypes are typically assumed to be used in cognitive processes that compute the similarity between a prototype and other representations, such as the representations of the objects to be categorized, in a linear manner. Exemplars are bodies of knowledge about individual members of a category (e.g., Kitty, Tiger), particular samples of a substance, and particular instances of a kind of event (e.g., my last visit to the physician). Exemplars are typically assumed to be used in cognitive processes that compute the similarity between a set of exemplars and other representations, such as the representations of the objects to be categorized, in a nonlinear manner.[31,32] Theories are bodies of causal, functional, generic, and nomological knowledge about categories, substances, types of events, etc.[33] Recent work on causal knowledge suggests that theories might be used in cognitive processes that are similar to the algorithms involved in causal reasoning.[34]

Recently, however, some philosophers of cognitive science have challenged the assumption that concepts share many properties.[10,11,35–41,h] They often put this point by asserting that concepts are not a natural kind, since in philosophy of science a natural kind is, roughly, a class of objects that non-accidentally share many relevant properties.[42] For these philosophers, there are a number of distinct kinds of concepts, which differ from one another in their important properties. On this view, it is erroneous to hope to develop a single theory of concepts that would apply to all kinds of concepts.

There are naturally different ways to develop this insight. Machery[10,35] and Weiskopf[39] have hypothesized that the class of concepts divides into exemplars, prototypes, and theories. On this view, each of the main theories of concepts turns out to be partly correct in that it correctly describes the important properties of a kind of concepts, while being incorrect in generalizing these properties to all concepts.

Let's consider Machery's hypothesis more closely (called 'the heterogeneity hypothesis'). According to this hypothesis, kinds (e.g., dogs), substances (e.g., water), and events (e.g., going to the dentist) are typically represented by several concepts. Thus, we have several concepts of dogs, several concepts of water, etc. These concepts belong to different kinds: for instance, we may have a prototype of dogs, a set of exemplars about particular dogs, and a theory of dogs (and similarly for other kinds, for substances, etc.). These types of concepts are typically used in distinct cognitive processes. For instance, Machery (Ref 10, chapter 6) argues that we have several distinct categorization processes, one using prototypes, one using exemplars, and one using theories (similarly for induction processes). Thus, it is typically the case that higher cognitive competences are underwritten by several distinct cognitive processes—an idea in the

spirit (though distinct from) dual-process theories of cognition (Ref 10, chapter 5).

Like Weiskopf,[10] Machery defends the hypothesis that the class of concepts includes exemplars, prototypes, and theories by noting that, when one examines 40 years of research on categorization and induction, one finds out that, in both areas of research, some phenomena are well explained if the concepts elicited by some experimental tasks are prototypes; some phenomena are well explained if the concepts elicited by other experimental tasks are exemplars; and yet other phenomena are well explained if the concepts elicited by yet other experimental tasks are theories. If one assumes that experimental conditions prime the reliance on one type of concepts (e.g., prototypes) instead of other types (e.g., exemplars and theories), this provides evidence for the heterogeneity hypothesis.

This argument can be usefully illustrated by considering the work on categorical induction—the capacity to conclude that the members of a category possess a property from the fact that the members of another category possess it and to evaluate the probability of this generalization (for review, see Refs 4, chapter 8; 5; and Ref 6). A large number of phenomena suggest that prototypes or exemplars are sometimes involved in induction.[43,44] Similarity-based models of induction, which assume that the processes underlying induction are defined over either prototypes or exemplars, explain best two well-known findings about induction—the similarity effect and the typicality effect. Other phenomena are best explained if the concepts involved in the relevant experimental conditions are causal theories. Investigating the judgments made by tree experts (landscapers, taxonomists, and park maintenance personnel) about the strength of inductive conclusions about trees, Proffitt et al.[45] found that, rather than relying on typicality (as predicted, for instance, by Osherson and colleagues' similarity-coverage model), the pattern of answers and the justifications provided suggest that experts often base their judgments on theories about hypothetical causal mechanisms (see also Ref 46). The fact that different properties of our inductive competence are best explained by theories positing different theoretical entities (viz., prototypes, exemplars, and theories) constitutes evidence for the existence of distinct kinds of concepts used in distinct processes. Strikingly, this conclusion is consistent with the emerging consensus among psychologists working on induction that people rely on several distinct induction processes.[4,5,45,47]

Recently, Machery and Seppälä[48] have attempted to provide some more direct evidence for the heterogeneity hypothesis. If this hypothesis is true, then some (and perhaps many) words express several distinct concepts: For instance, the predicate 'vegetable' may express a prototype of vegetables (which would represent the typical properties of vegetables, e.g., that they are used in salad, that they are typically salty, not sweet) and a theory of vegetables (which would perhaps represent vegetables as having necessary properties). If this is correct, then people should be willing to say that something is an x and that it is not an x if it is classified as an x when $\ulcorner x \urcorner$ expresses one kind of concept (e.g., a prototype) and not classified as an x when it expresses another kind of concept (e.g., a theory). For instance, people should be willing to say that tomatoes *are* and *are not* vegetables when, respectively, 'vegetables' express the prototype of vegetables and the theory of vegetables. Machery and Seppälä[48] present some preliminary data that support this hypothesis. Participants were presented with pairs of sentences (Table 1), and they were asked whether each sentence was true on a 7-point scale (see Ref 48 for further detail). The proportion of participants agreeing with both sentences was measured. It was found that for all pairs the proportion of participants agreeing was higher than for control pairs. Furthermore, while the proportion of agreement to both sentences varied substantially across pairs, there was always at least a minority of participants who agreed with the two sentences. For instance, about a quarter of participants agreed with both sentences of pair C.[i] While more research is needed to confirm and extend these findings, the heterogeneity hypothesis predicts this pattern of responses.

TABLE 1 | Stimuli in Ref 48

Pair	First Sentence	Second Sentence
A	In a sense, tomatoes are vegetables	In a sense, tomatoes are not vegetables
B	In a sense, penguins are birds	In a sense, penguins are not birds
C	In a sense, whales are fish	In a sense, whales are not fish
D	In a sense, a piano is a piece of furniture	In a sense, a piano is not a piece of furniture
E	In a sense, chess is a sport	In a sense, chess is not a sport
F	In a sense, zombies are alive	In a sense, zombies are not alive

Let us now turn to other versions of the insight that the class of concepts divides into different kinds. Highlighting the distinctive role of language in human cognition and endorsing a dual-process cognitive architecture, Piccinini has argued that concepts divide

into two kinds—explicit and implicit concepts (see also Refs 37 and 49). According to Piccinini, explicit concepts are used in processes that are similar to what psychologists call System 2—viz. processes that are controlled, linguistic, and accessible to conscious inspection. By contrast, implicit concepts are used in processes that are similar to what psychologists call System 1—viz. processes that are automatic, nonlinguistic, and non-accessible to conscious inspection. Implicit concepts do not depend on the possession of a language, and are found in nonlinguistic creatures, while the possession of explicit concepts supposes the possession of a language.

Piccinini's and others' suggestions are intriguing. It is a common view that the emergence of language during the evolution of hominids has massively modified cognition, and dual-process cognitive architectures are popular in cognitive science. Nonetheless, Machery[36,50] has argued against these suggestions. First, Machery highlights the evidence that concepts of unlexicalized categories, which are plausibly classified as implicit concepts in Piccinini's scheme as they are not associated with any word and as their content is not consciously accessible, and concepts used in linguistic categorization and induction have similar properties. This finding speaks against the idea that implicit and explicit concepts form two very different kinds of concepts. Second, little evidence supports the speculation that the evolution of language in the human lineage resulted in a fundamental modification of our conceptual repertoire. Obviously, language does have an impact on which concepts people have because people acquire many concepts by talking to other people and because linguistic communication also enables cultures to develop and to transmit to their members concepts that would have not been developed without linguistic communication. But acknowledging this does not amount to recognizing that the emergence of language in the hominid lineage resulted in the emergence of a new, distinct kind of concepts. Third, while this is not the place to develop this concern at length, Machery (Refs 10, chapter 5; 50) cautions against embracing dual-process theories, which are so far formulated so vaguely that they tend to accommodate empirical findings instead of predicting them.

Philosophers of cognitive science who have highlighted the diversity of the kinds of concepts have taken two opposite positions about the usefulness of the theoretical term 'concept' in cognitive science. Machery[10,11,35] has argued that in science the failure of a theoretical term to pick out a natural kind gives scientists a defeasible reason to abandon this term. This reason is defeated when this term plays a useful role in a scientific scheme despite failing to pick out a natural kind. If the arguments discussed in this section are on the right track, the theoretical term 'concept' fails to pick out a natural kind, and, Machery argues, maintaining this term within the theoretical scheme of cognitive science would lead to more costs than benefits (for a critical discussion of this argument, see Refs 51–54). Hence, he recommends that 'concept' be eliminated from the theoretical vocabulary of cognitive science and replaced with theoretical terms such as 'exemplar', 'prototype', and 'theory', which refer to the bodies of knowledge actually used in the processes underlying our cognitive competences.

Some philosophers and cognitive scientists take a different line, opposing various forms of pluralism to Machery's eliminativism. Pluralism about concepts acknowledges that the class of concepts divides into kinds that are very different from one another, but maintains that in spite of these differences the term 'concept' remains scientifically useful (see, particularly, Refs 8, 39, 55, 56, and 57). Weiskopf[39] maintains that some properties are shared by all concepts and that eliminating concepts would prevent scientists from studying these shared properties. In a similar vein, Danks[55] shows that various models of categorization can be represented as graphical models, and concludes that there are abstract properties shared by all these models. Finally, Weiskopf[57] holds that there is a scientifically important difference between the organisms that have concepts and those that do not, and that the notion of concept is needed in cognitive science to draw this distinction.

NEO-EMPIRICISM

A traditional view in cognitive science holds that concepts are amodal representations—their format is distinct from that of the representations used in the sensory, motor, and affective systems. It is also often held that amodal representations are linguistic, but amodal representations need not form a language. Recently, however, a number of psychologists and philosophers have argued against this view in favor of concept empiricism.[2,58–63] The basic idea behind this movement, which is gaining traction in the cognitive sciences, is that concepts are composed of perceptual and motor representations and that thinking (e.g., categorizing objects or drawing inferences) involves deploying these representations in episodes of perceptual simulation. On this view, the concept HAMMER would consist of a set of perceptual representations (what a hammer looks like, how it feels to hold and swing one, etc.), and entertaining this concept would consist in simulating seeing a hammer, holding it, etc. Following Machery,[64] we refer to theorists who claim

that all concepts are composed of perceptual representations as 'neo-empiricists'.

Neo-empiricism is a revival of an old idea, drawing its core tenets from the British empiricists of the 17th and 18th centuries. These philosophers, such as John Locke and David Hume, argued that all concepts are impressions (mental traces of past sensory experiences) or copies of impressions. Neo-empiricists similarly hold that concepts are inherently perceptual, but they differ from traditional concept empiricists in several aspects.[2,59,60] First, for neo-empiricists, concepts need not be conscious. Second, neo-empiricists are not necessarily antinativist as they disentangle the claim that concepts are perceptual from the thesis that there are no innate concepts. Third, neo-empiricists embrace a much broader definition of 'perceptual' than traditional concept empiricists: For the neo-empiricist, 'perceptual' encompasses a variety of sensory, motor, affective, and other introspective (e.g., proprioception) capacities.

According to neo-empiricists, their theory has two distinct advantages over amodal accounts of conceptual knowledge—it is more parsimonious than amodal theories (since sensory, motor, and affective representations do all the work that amodal symbols are thought to do), and it does a better job in explaining how cognitive processes interface with perception and behavior. In addition to these theoretical considerations, neo-empiricists have furnished a large body of behavioral and neuroscientific data in support of the view that concepts are perceptual. In the remainder of this section, we focus on this body of data.

Drawing from cognitive psychology, neo-empiricists appeal to a variety of sensorimotor effects on reaction times and other behavioral measures in property verification tasks, property listing tasks, and other tasks to bolster the claim that entertaining a concept consists in engaging in a perceptual simulation.[65–71]

In a representative study, Pecher et al.[72] presented participants with a series of pairs of nouns and predicates in a property verification task—subjects had to judge whether the predicate is true of the objects denoted by the noun. In each pair (e.g., BLENDER-loud and LEAVES-rustling or BLENDER-loud and CRANBERRIES-tart), the first property (e.g., loud) was followed with a property in either the same (e.g., rustling) or a different (tart) sensory modality. Pecher et al. found that participants were faster at making a judgment about the second member of a pair when it involved the same modality (e.g., loud and rustling) as the first and slower when the pair involved two different modalities (e.g., loud and tart). This suggests that switching modalities during the task results in a cost in processing speed. Pecher et al. argue that amodal theories do not predict such a switching cost, and take these findings to support neo-empiricism.

Cognitive neuroscience [particularly via functional magnetic resonance imaging (fMRI) and other forms of neuroimaging] provides another strand of research frequently used to support neo-empiricism. Neo-empiricists often cite findings that conceptual processing activates perceptual (sensory and motor) brain regions. Neuroscientists have found, for example, that reading certain verbs (e.g., 'grasp' or 'kick') activates areas of premotor cortex involved in voluntary movement of the body part (in this case, hand or foot) corresponding to the appropriate action.[73] The neo-empiricist takes this and other results to show that comprehending verbs involves simulating the actions they denote.

Neo-empiricists have also appealed to neuroscience to support the thesis that in different semantic domains—particularly, animals versus tools—concepts tend to depend on different perceptual capacities because different perceptual capacities are important when one interacts with, e.g., animals and tools. Sim and Kiefer[74] reported that visual areas of the cortex are preferentially active when participants think about animals, while regions of the motor cortex are engaged when participants think about artifacts.

Amodal theorists have raised a number of objections against neo-empiricism.[64,75–79] Machery[75,76] has recently identified three methodological issues facing the research in support of neo-empiricism. The first difficulty, which Machery calls 'Anderson's problem' in reference to Anderson's[80] contribution to the debate between propositional and imagistic accounts of thinking, concerns the fact that neo-empiricists claim that conceptual processing exhibits sensorimotor effects (e.g., modality switching costs) that are not predicted by the amodal approach to concepts. Machery highlights that some amodal theories make the same predictions as neo-empiricism regarding sensorimotor effects. Louwerse and Connell,[81] for example, point out that the modality switching costs observed by Pecher et al.[72] could arise from the connections between a set of amodal representations. According to this view, statistical co-occurrence might dictate that a reader who encounters the word 'loud', may also activate words such as 'baby' or 'blender' without any involvement of the auditory system. In short, linguistic context may give a set of amodal representations structure (e.g., one in which switching between clusters of less frequently co-occuring concepts would result in a processing cost) resembling that of a semantic system whose content is organized by modality. Anderson's problem demonstrates that neo-empiricist

findings cannot disconfirm the amodal approach in general; they can only provide evidence for and against specific amodal theories.

Machery's second objection is the 'problem from imagery'. Typically, amodal theorists do not claim that individuals never use perceptual simulation to solve cognitive tasks. For example, visualizing walking through one's home might be a viable strategy when asked, 'What color is your bedroom door?' Amodal theorists merely pose that, in addition to mental imagery, people also possess amodal concepts. Briefly put, the problem from imagery states that showing that people use imagery in some cognitive tasks falls short of providing evidence for neo-empiricism and against amodal theories of concepts when one would *a priori* expect people to solve this task by means of imagery. Even worse, Machery[64] argues that evidence from cognitive psychology supporting neo-empiricism may largely derive from an idiosyncratic subset of tasks that prime participants to employ mental imagery (but see Ref 82 for a response).

Some prominent neo-empiricists have responded to Anderson's problem and the problem from imagery.[60,82] Many neo-empiricists endorse Barsalou's[60] contention that there is little empirical evidence for amodal representations. In response to Anderson's problem, Prinz[82] draws on this alleged lack of support for amodal symbols to argue that, once amodal theorists admit that people often use perceptual simulations in cognitive tasks, they become vulnerable to an argument from parsimony: Even if some amodal models can predict neo-empiricist findings, these models are less parsimonious. Furthermore, in response to the problem from imagery, Prinz notes that individuals seem to use perceptual imagery to identify what properties are typical of certain categories (see Ref 71 for a discussion of the findings). According to Prinz, using amodal symbols to retrieve such common, overlearned features of categories would be faster than using perceptual simulations. So, we would not expect *a priori* that people would use perceptual representations to solve this task.

There are two problems with Prinz's argument from parsimony. First, it does not assuage the worry that evidence of sensorimotor effects in conceptual processing simply cannot speak for or against amodal theories in general since parsimony, not the empirical evidence, is meant to undermine these theories. Second, Prinz's argument is substantially weakened if there is compelling evidence for the existence of amodal representations. Machery,[64] Dove,[78] and others (e.g., Hauser and Spelke[83]) have convincingly argued that number approximation, which exhibits a striking degree of intermodal transfer and occupies a set of brain regions seemingly distinct from any perceptual system, involves amodal representations. With such evidence for amodal representations, there is little reason to prefer neo-empiricist theories over amodal ones on parsimony grounds, and both Anderson's problem and the problem from imagery refuse to dissolve.

Evidence for amodal representations leads to Machery's third problem for neo-empiricism: the so-called generality problem. The generality problem holds that, even if neo-empiricists successfully show that some kinds of concepts are perceptual representations, this fails to establish that all concepts are perceptual. In line with Machery's[10] heterogeneity hypothesis, perceptual representations may turn out be *one kind* of concept, with amodal representations forming other types of concepts. Thus, for Dove,[78] a variety of empirical findings such as imageability effects and the existence of semantic deficits related to abstract, but not concrete, words (Refs 84 and 85, respectively) suggest that concrete concepts are stored as perceptual representations, while more abstract concepts rely on a distinct representational system.

Amodal theorists have also challenged the case for neo-empiricism from cognitive neuroscience.[11,77,86] Neo-empiricists claim that entertaining a concept involves our perceptual (sensory and motor) brain areas. Amodal theorists have raised two objections against this claim: (1) conceptual processing seems to occur, at least sometimes, in brain regions distinct from any perceptual system and (2) activations that are observed in perceptual regions in conjunction with conceptual processing may be downstream effects rather than reflect in which areas concepts are retrieved and processed. We elaborate on these two objections below.

While neo-empiricists predict that perceptual and conceptual processing will share the same neural substrate, a typical finding in cognitive neuroscience[87] is that conceptual processing activates areas close to, but not identical to, perceptual regions. Willems et al.,[88] for example, found that comprehending certain verbs (e.g., 'to kick'), performing the relevant action (kicking), and simulating this action (imagining kicking) elicited activity in neighboring, but distinguishable regions of premotor cortex.

Furthermore, several studies suggest that patterns of brain activation found during conceptual processing do not depend on the perceptual properties of what concepts refer to. Bedny et al.[86] found that portions of the posterior lateral temporal cortex (PLTC) responded equally to verbs that referred to events with significantly different visual-motion properties: PLTC activation was equivalent for action

verbs, such as 'to run', and mental verbs, such as 'to think'. The authors conclude that activation in the PLTC may reflect aspects of conceptual processing abstracted away from sensorimotor experience, such as the retrieval of event concepts or grammatical information pertaining to verbs.

Mahon and Caramazza[77] and Machery[76] have argued that fMRI lacks the temporal resolution to determine whether the activation in perceptual regions during conceptual retrieval reflects conceptual processing. During concept retrieval patterns of activation could rapidly spread from the areas in which concepts are retrieved and processed to perceptual regions. Because fMRI aggregates activity over a broad timing window (~1–5 s), the data taken to support neo-empiricism collapse BOLD activity reflecting conceptual processing and spreading activation into a single signal. A number of neo-empiricists Boulenger et al.,[89] Hoenig et al.[90] have responded that the activation (e.g., measured with event-related potentials) in perceptual regions (e.g., motor cortex) that correlates with semantic task performance takes place too quickly (~100–200 ms after stimulus onset) to constitute post-conceptual processing. However, without offering a principled comparison of different models of meaning access (accounts of where concepts are stored and how long it takes to deploy them), this kind of rapid activation may still reflect post-conceptual or even (e.g., in cases where meaning access and behavioral preparation occur in parallel after a lexical stimulus) para-conceptual processing.[77]

While neo-empiricists have recently argued that neuropsychology supports the thesis that concepts are composed of perceptual representations, recent work on semantic dementia, a neurodegenerative disorder causing selective semantic impairments, undermines their argument. Neo-empiricists sometimes point to category-specific semantic deficits resulting from damage to specific perceptual systems.[60] Particularly, damage to motor cortex is frequently correlated with deficits in knowledge about artifacts, while lesions in the visual system are more likely to impair knowledge of living things.[91,92] For neo-empiricists, this is evidence that to entertain a concept consists in engaging in perceptual simulation, which requires intact perceptual brain areas.

However, neuropsychology—particularly, research on semantic dementia—is not particularly supportive of neo-empiricism. Semantic dementia is a form of frontotemporal dementia characterized by bilateral degeneration of tissue around the anterior temporal lobe (ATL) and a progressive loss of conceptual knowledge.[93] Patients with semantic dementia experience debilitating impairments in the ability to recognize, name, and access relevant information about objects. Two symptoms of semantic dementia are problematic for the neo-empiricist picture. First, damage to a single anatomical locus (the ATL) results in a loss of conceptual knowledge across the broad domains for which other patients can acquire specific deficits (e.g., patients will lose knowledge about *some* animals, *some* tools, etc., while others are spared). Because neo-empiricists seem to be committed to the claim that the conceptual system is distributed over several modal systems throughout the brain, this finding is difficult to explain for neo-empiricism. Second, semantic dementia patients exhibit highly multimodal deficits related to specific affected concepts.[94] Patients lose all feature knowledge (visual, auditory, tactile, etc.) for specific concepts, while knowledge about those same kinds of features for related concepts remains intact.[95] This pattern of deficits, which we call 'modality-general, item-specific', is challenging for neo-empiricism, which predicts that people would lose the capacity to represent the features in a given modality (e.g., all visual features) across all concepts, while the capacity for representing the features in other modalities would be preserved (a 'modality-specific, item-general' pattern).

Neo-empiricism is a provocative thesis that has spurred a wealth of productive research into the nature of concepts and conceptual processing. However, in light of the objections raised by amodal theorists, we urge greater caution in interpreting the alleged findings for neo-empiricism from psychology, cognitive neuroscience, and neuropsychology than is typically found in the literature. In light of our discussion, it is possible that all concepts are amodal representations or that perceptual simulations are merely one kind of concept.

CONCEPT LEARNING

Humans have an almost limitless capacity to acquire new concepts (think about NEUTRINO or BLOGOSPHERE). Considering the generative ability of the adult conceptual system invites the question: Where do our concepts come from?

Many concepts arise from combinatorial processes that use existing concepts to form novel ones (e.g., an ATM is a MACHINE that is a BANK TELLER; on concept combination, see many chapters in Ref 96). If new concepts arise through combination, there must be a stock of unstructured or primitive concepts from which structured concepts derive.[97–101] So, where do our primitive concepts come from?

A natural response is that primitive concepts are innate, while structured ones are learned. However,

it is extremely unlikely that a small number of innate concepts (such as CAUSE, AGENT, and other purportedly innate concepts) can be combined to form the entire suite of concepts in the adult conceptual system. Therefore, it seems that we must either possess an enormous stock of innate concepts or a kind of learning that generates new primitive concepts.

This observation raises two distinct, but related issues about concept acquisition. First, while most theorists[101] think we have an innate repertoire of primitive concepts, there is wide disagreement over the nature (domain-specific vs domain-general, perceptual vs non-perceptual, etc.) and extent of this innate repertoire. Second, there are worries about whether it is even possible to learn new primitive concepts[100,102] and, if it is, what processes are needed to learn them. To illustrate why the notion of learning new primitive concepts is problematic, we turn to Fodor's[97,98] argument for radical concept nativism.

Fodor[97,98] famously argued for radical concept nativism, the proposal that almost our entire stock of lexical concepts (concepts such as RADISH or UMBRELLA that correspond to natural language words) is innate. The thrust of the argument is that lexical concepts are primitive—viz., they lack structure—and primitive concepts cannot be learned. Fodor begins with the uncontroversial premise that (barring bizarre cases in which swallowing pills, bumps on the head, or other nondevelopmental, non-psychological processes generate new concepts) concepts are either learned or innate. His second premise is that, while structured concepts are simply learned by combining primitives, primitive concepts must be learned through hypothesis testing (Ref 97, p. 80–82). His third premise is that hypothesis testing requires possession of the target concept: To formulate a hypothesis about a concept, one needs to possess this concept in the first place. As a consequence, hypothesis testing fails to be a kind of learning, and therefore primitive concepts cannot be learned.

To get a sense of Fodor's claim that hypothesis testing presupposes having the target concept, imagine teaching a learner the concept THREE by showing her cards with varied numbers of shapes and providing feedback on whether she correctly applies 'three' to the cards.[j] Now suppose that the learner makes the following series of designations:

1. Applies 'three' to a card with three circles and receives positive feedback.
2. Applies 'three' to a card with one circle and receives negative feedback.
3. Subsequently, applies 'three' to all cards with any three shapes.

This seems like a canonical example of learning a concept. However, Fodor points out that the progression from 1 to 3 requires the learner to entertain two hypotheses—H_1: 'three' refers to circles or H_2: 'three' refers to three—and subsequently rule out H_1. The problem, according to Fodor, is that entertaining H_1 and H_2 presupposes that one is able to think about circles and three, which requires the very concepts one is supposed to be learning.

Fodor also claims that lexical concepts are primitive (see the discussion of atomism in the section on *What Are Concepts*?), and concludes that, as they cannot be learned, our entire stock of lexical concepts must be innate.[k] However, Fodor's argument for radical concept nativism does not require that lexical concepts (e.g., REFRIGERATOR) are primitive—even if typical lexical concepts have structure, we will still possess thousands of extremely counterintuitive primitive concepts.[l] The upshot of Fodor's argument is that there can be no psychological explanation—viz., one that appeals to learning—of primitive concept acquisition.[99] Instead, Fodor relegates primitive concept acquisition to the realms of biological (1981) or metaphysical (1998) explanation.[103]

Most philosophers and psychologists reject radical concept nativism, which seems to fly in the face of common sense or clash with extant empirical data on concept acquisition.[102,104] Churchland[104] claims that radical concept nativism presents a *reductio ad absurdum* of Fodor's Language of Thought framework, while Putnam[105] argues against radical concept nativism on the grounds that natural selection could not have produced such an innately rich conceptual system. However, these responses miss the mark because they merely reject Fodor's radical conclusion without explaining where he went wrong.[99]

Several theorists have attempted to blunt the strength of Fodor's argument for radical concept nativism by challenging the premise that learning requires hypothesis testing.[99,100,102,106,107] Samet and Flanagan,[106] for example, argue that numerous forms of learning (including food aversion and motor learning) occur implicitly or in a single trial, and therefore do not involve hypothesis testing. Fodor would likely reply that episodes of food aversion do not count as genuine instances of, and are not good models for, concept learning. Furthermore, the notion that hypothesis testing is the only form of conceptual learning is not crucial to Fodor's point that concepts without structure cannot be learned. Any satisfying answer to Fodor needs to directly address the challenge of how we acquire new primitive concepts.

So, how can primitive concepts be learned? Carey[7,101] has recently presented a novel account of

the origin of concepts (see Refs 99 and 100 for alternative accounts). According to Carey,[101] the conceptual system begins with an innate set conceptual primitives—which she calls 'core cognition'—that subserve the domains of causal and spatial reasoning, agency, and representations of number. In addition to core cognition, which remains largely unchanged throughout an individual's lifetime, the conceptual system has a learning process that generates new conceptual primitives.

Carey[7,101] argues that new conceptual primitives arise through 'Quinean bootstrapping', a learning process that allegedly generates new concepts in the absence of conceptual resources needed to form hypotheses about target concepts.[m] In episodes of Quinean bootstrapping, mental symbols are created as placeholders, and they derive their meaning from their relations to other symbols. The meaning of the placeholders is initially determined by existing conceptual structure, but, through processes such as analogy, limiting case analysis, and thought experiment, coined symbolic representations can eventually take on new primitive meanings. For example, Carey (Ref 101, p. 121–122) argues that children acquire the concepts WEIGHT and DENSITY by using a set of placeholder representations (e.g., the equation 'density = weight/volume' and box-dot diagrams). While the meaning of these placeholders is first determined by their relations to preexisting concepts (e.g., FELT WEIGHT), they take on new primitive meanings as learners map these representations to measurable physical quantities. According to Carey, Quinean bootstrapping enables learners to acquire new primitive concepts, such as the concept of natural number, by reasoning with placeholders.

Unfortunately, Quinean bootstrapping does not seem to solve Fodor's puzzle of concept acquisition. Consider how children are hypothesized to acquire the concept of natural number (Ref 7, chapters 8 and 9). Initially, they know the counting routine 'one, two, three, four', and each numeral in this routine acts as a placeholder for the relevant number concept (ONE, TWO, etc.). Then, each word in this list becomes associated with: (1) a mental model of a set ('one' is associated with a mental model with one element, 'two' with a mental model with two elements, etc.) and (2) a procedure that allows children to apply each word to its corresponding set (viz., children know that 'one' applies to sets composed of a single individual, etc.). These four mental models form a sequence, and, when this sequence is put in parallel with the counting routine, children realize that each numeral refers to a set that has exactly one individual more than the set referred to by the previous numeral in the counting list. At this point, 'one' genuinely expresses ONE, 'two' genuinely expresses TWO, etc., and children are able to understand natural numbers beyond four.

While this bootstrapping hypothesis is ingenious, there are many different ways of ordering the relevant mental models into a sequence, and it is unclear why children should infer that each mental model in the sequence is obtained by adding one individual to the previous mental model in the sequence rather than any of the possible ways of ordering these mental models. Furthermore, one may wonder whether bootstrapping is really different from hypothesis testing: Shouldn't the child be able to form the hypothesis that each mental model in the sequence is obtained by adding one individual to the previous mental model in the sequence, which supposes the possession of the concept of addition and thus the concept of number? It appears that Fodor's problem of concept acquisition remains a puzzle for philosophers and psychologists to solve.

NOTES

[a] Space limits prevented us from examining a few important controversies, such as the relationship between conceptual and linguistic systems, including the Sapir-Whorf hypothesis, the variation of concepts across cultures, and the nature of animal concepts.

[b] See, e.g., Refs 13, 16, 17, 108, and 109; for review see Refs 110 and 111.

[c] For review, see Refs 4, 10, and 28.

[d] Of course, we do not mean to deny that some philosophers of cognitive science are also[2] or sometimes primarily[10] interested in the issues examined by the theories of concepts developed in cognitive science.

[e] We will use small caps to name concepts.

[f] See Ref 18 for a critical discussion of this proposal.

[g] Malt[112] has criticized this characterization on the grounds that there are no such things as bodies of knowledge used *by default* in the cognitive processes underlying our capacity to categorize, draw inductions, etc. (see Refs 50 and 113 for a response). Slaney and Racine[114] have criticized it on the grounds that it is committed to a philosophically suspect mentalism (see Ref 115 for a response).

[h] For critical discussion, see, e.g., Refs 55, 116, 117–119.

[i] One could justifiably wonder whether the heterogeneity hypothesis is undermined by the fact that 75% of participants did not agree with both sentences. However, it is worth keeping in mind that, even if the heterogeneity hypothesis is true, many people may associate words with one of several existing concepts

(e.g., 'fish' with a theory of fish instead of a prototype of fish). This would explain why for many people many words are not polysemous.

[j]We owe this example to Ref 99.

[k]As Laurence and Margolis[99] note, this amounts to claiming that individuals possess significantly more than 60,000 innate concepts (by estimating average natural language vocabularies).

[l]To see why this is the case, assume that the lexical concept ROSE is composed of the concepts PRETTY, RED, and FLOWER. Once we find a constituent concept that lacks structure (e.g., PRETTY), we will have to concede that it is primitive and therefore cannot be learned. This will still yield a form of radical concept nativism.

[m]Quinean bootstrapping has also been hypothesized to occur during episodes of scientific discovery, in which researchers create concepts for newly discovered phenomena that are initially understood in terms of their existing theories.[120]

REFERENCES

1. Fodor JA. Is it a bird? Problems with old and new approaches to the theory of concepts. *Times Lit Suppl* 2003, 17:3–4.
2. Prinz JJ. *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, MA: MIT Press; 2002.
3. Osherson DN, Smith EE, Wilkie O, López A, Shafir E. Category-based induction. *Cognit Sci* 1990, 15:251–269.
4. Murphy G. *The Big Book of Concepts*. Cambridge, MA: NIT Press; 2002.
5. Sloman SA, Lagnado D. The problem of induction. In: Holyoak K, Morrison R, eds. *The Cambridge Handbook of Thinking & Reasoning*. New York: Cambridge University Press; 2005, 95–116.
6. Feeney A, Heit E, eds. *Inductive Reasoning: Experimental, Developmental, and Computational Approaches*. Cambridge: Cambridge University Press; 2007.
7. Carey S. *The Origin of Concepts*. New York: Oxford University Press; 2009.
8. Lombrozo T. The campaign for concepts. *Dialogue* 2011, 50:165–177.
9. Margolis E, Laurence S. Concepts and theoretical unification. *Behav Brain Sci* 2010, 33:219–220.
10. Machery E. *Doing Without Concepts*. New York: Oxford University Press; 2009.
11. Machery E. The heterogeneity of knowledge representation and the elimination of *concept*. *Behav Brain Sci* 2010, 33:231–244.
12. Block N. Advertisement for a semantics for psychology. *Midwest Stud Philos* 1986, 10:615–678.
13. Peacocke C. *A Study of Concepts*. Cambridge, MA: MIT Press; 1992.
14. Block N. Advertisement for a semantics for psychology. *Midwest Studies in Philosophy* 1986, 10:615–678.
15. Fodor JA. Concepts: a pot-boiler. *Cognition* 1994, 50:95–113.
16. Fodor JA. *Concepts, Where Cognitive Science Went Wrong*. New York: Oxford University Press; 1998.
17. Fodor JA. *LOT 2: The Language of Thought Revisited*. New York: Oxford University Press; 2008.
18. Fodor JA, Lepore E. *Holism: A Shopper's Guide*. Cambridge, MA: Blackwell; 1992.
19. Fodor JA, Lepore E. *The Compositionality Papers*. New York: Oxford University Press; 2002.
20. Prinz JJ. Regaining composure: a defense of prototype compositionality. In: Werning M, Hinzen W, Machery E, eds. *The Oxford Handbook of Compositionality*. Oxford: Oxford University Press. In press.
21. Rey G. The analytic/synthetic distinction. In: Zalta EN, ed. The Stanford Encyclopedia of Philosophy (Winter 2010 Edition). Available at: http://plato.stanford.edu/archives/win2010/entries/analytic-synthetic/.
22. Quine WVO. Two dogmas of empiricism. *Philos Rev* 1951, 60:20–43.
23. Grice P, Strawson P. In defense of a dogma. *Philos Rev* 1956, 65:141–158.
24. Boghossian PA. Analyticity reconsidered. *Nous* 1996, 30:360–391.
25. Johnson EK. From impossible words to conceptual structure: the role of structure and processes in the lexicon. *Mind Lang* 2004, 19:334–358.
26. Russell G. *Truth in Virtue of Meaning: A Defence of the Analytic/Synthetic Distinction*. New York: Oxford University Press; 2008.
27. Machery E. 100 years of psychology of concepts: the theoretical notion of concept and its operationalization. *Stud Hist Philos Biol Biomed Sci* 2007, 38:63–84.
28. Machery E. Concepts: a tutorial. In: Belohlavek R, Klir G, eds. *Concepts and Fuzzy Logic*. Cambridge, MA: MIT Press; 2011, 13–44.
29. Rosch E, Mervis CB. Family resemblance: studies in the internal structure of categories. *Cognit Psychol* 1975, 7:573–605.

30. Hampton JA. Polymorphous concepts in semantic memory. *J Verbal Learn Verbal Behav* 1979, 18:441–461.
31. Medin DL, Schaffer MM. Context theory of classification learning. *Psychol Rev* 1978, 85:207–238.
32. Nosofsky RM. Attention, similarity, and the identification-categorization relationship. *J ExpPsychol Learn Mem Cogn* 1986, 10:104–114.
33. Murphy GL, Medin DL. The role of theories in conceptual coherence. *Psychol Rev* 1985, 92:289–316.
34. Gopnik A, Glymour C, Sobel D, Schulz L, Kushnir T, Danks D. A theory of causal learning in children: Causal maps and Bayes nets. *Psychol Rev* 2004, 111:1–31.
35. Machery E. Concepts are not a natural kind. *Philos Sci* 2005, 72:444–467.
36. Machery E. How to split concepts. Reply to Piccinini and Scott. *Philos Sci* 2006a, 73:410–418.
37. Piccinini G, Scott S. Splitting concepts. *Philos Sci* 2006, 73:390–409.
38. Frixione M. Do concepts exist? A naturalistic point of view. In: Penco C, Beaney M, Vignolo M, eds. *Explaining the Mental. Naturalist and Non-Naturalist Approaches to Mental Acts and Processes*. Cambridge: Cambridge Scholars Publishing; 2007.
39. Weiskopf DA. The plurality of concepts. *Synthese* 2009, 169:145–173.
40. Piccinini G. Two kinds of concepts: implicit and explicit. *Dialogue* 2011, 50:179–193.
41. Park JJ. (Forthcoming). Prototypes, exemplars, and theoretical and applied ethics. *Neuroethics*.
42. Boyd R. Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philos Stud* 1991, 61:127–148.
43. Osherson DN, Smith EE, Wilkie O, Lopez A, Shafir E. Category-based induction. *Psychol Rev* 1990, 97:185–200.
44. Sloman SA. Feature-based induction. *Cognit Psychol* 1993, 25:231–280.
45. Proffitt JB, Coley JD, Medin DL. Expertise and category-based induction. *J Exp Psychol Learn Mem Cogn* 2000, 26:811–828.
46. López A, Atran S, Coley JD, Medin DL, Smith EE. The tree of life: universal and cultural features of folkbiological taxonomies and inductions. *Cognit Psychol* 1997, 32:251–295.
47. Rehder B. When causality and similarity compete in category-based property induction. *Mem Cogn* 2006, 34:3–16.
48. Machery E, Seppälä S. Against hybrid theories of concepts. *Anthropol Philos* 2010, 10:97–130.
49. Poirier L, Beaulac G. Le véritable retour des définitions. *Dialogue* 2011, 50:153–164.
50. Machery E. Replies to Lombrozo, Piccinini, and Poirier and Beaulac. *Dialogue* 2011, 50:195–212.
51. Edwards K. Unity amidst heterogeneity in theories of concepts. *Behav Brain Sci* 2010, 33:210–211.
52. Edwards K. Higher-Level concepts and their heterogeneous implementations: a polemical review of Edouard Machery's *Doing Without Concepts*. *Philos Psychol* 2011, 24:119–133.
53. Gonnerman C, Weinberg J. Two uneliminated uses for "concepts": hybrids and guides for inquiry. *Behav Brain Sci* 2010, 33:211–212.
54. Samuels R, Ferreira M. Why don't concepts constitute a natural kind? *Behav Brain Sci* 2010, 33:222–223.
55. Danks D. Theory unification and graphical models in human categorization. In: Gopnik A, Schulz L, eds. *Causal Learning: Psychology, Philosophy, and Computation*. New York: Oxford University Press; 2007, 173–189.
56. Hampton JA. Concept talk cannot be avoided. *Behav Brain Sci* 2010, 33:212–213.
57. Weiskopf DA. The theoretical indispensability of concepts. *Behav Brain Sci* 2010, 33:228–229.
58. Glenberg AM. What memory is for. *Behav Brain Sci* 1997, 20:1–55.
59. Barsalou LW. Perceptual symbol systems. *Behav Brain Sci* 1999, 22:577–660.
60. Barsalou LW. Grounded cognition. *Annu Rev Psychol* 2008, 59:617–645.
61. Gallese V, Lakoff G. The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cogn Neuropsychol* 2005, 22:455–479.
62. Pecher D, Zwaan RA. *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*. Cambridge: Cambridge University Press; 2005.
63. Prinz JJ. The return of concept empiricism. In: Cohen H, Lefebvre C, eds. *Categorization and Cognitive Science*. Amsterdam: Elsevier; 2005, 679–695.
64. Machery E. Concept empiricism: a methodological critique. *Cognition* 2007, 104:19–46.
65. Borghi AM, Glenberg AM, Kaschak MP. Putting words in perspective. *Mem Cogn* 2004, 32:863–873.
66. Meteyard L, Bahrami B, Vigliocco G. Motion detection and motion verbs: language affects low-level visual perception. *Psychol Sci* 2007, 18: 1007–1013.
67. Pecher D, Van Dantzig S, Schifferstein NJ. Concepts are not represented by conscious imagery. *Psychon Bull Rev* 2009, 5:914–919.
68. Solomon KO, Barsalou LW. Representing properties locally. *Cognit Psychol* 2001, 43:129–169.
69. Solomon KO, Barsalou LW. Perceptual simulation in property verification. *Mem Cogn* 2004, 32:244–259.

70. Vermeulen N, Corneille O, Niedenthal PM. Sensory load incurs conceptual processing costs. *Cognition* 2008, 109:287–294.
71. Wu L. *Perceptual Representation in Conceptual Combination*. PhD thesis, University of Chicago, Chicago, IL, 1995.
72. Pecher D, Zeelenberg R, Barsalou LW. Verifying properties from different modalities for concepts produces switching costs. *Psychol Sci* 2003, 14: 119–124.
73. Hauk O, Johnsrude I, Pulvermüller F. Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 2004, 41:301–307.
74. Sim EJ, Kiefer M. Category-related brain activity to natural categories is associated with the retrieval of visual features: Evidence from repetition effects during visual and functional judgments. *Cogn Brain Res* 2005, 24:260–273.
75. Machery E. Two dogmas of neo-empiricism. *Philos Compass* 2006, 1:398–412.
76. Machery E. Reply to Barbara Malt and Jesse Prinz. *Mind Lang* 2010, 25:634–646.
77. Mahon BZ, Caramazza A. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J Physiol Paris* 2008, 102:59–70.
78. Dove G. Beyond perceptual symbols: a call for representational pluralism. *Cognition* 2009, 110:412–431.
79. Dove G. On the need for embodied and dis-embodied cognition. *Front Psychol* 2011, 1:1–13.
80. Anderson JR. Arguments concerning representations for mental imagery. *Psychol Rev* 1978, 85:249–277.
81. Louwerse M, Connell L. A taste of words: linguistic context and perceptual simulation predict the modality of words. *Cogn Sci* 2011, 35:381–398.
82. Prinz JJ. Can concept empiricism forestall eliminativism? *Mind Lang* 2010, 25:612–621.
83. Hauser MD, Spelke ES. Evolutionary and developmental foundations of human knowledge: a case study of mathematics. In: Gazzaniga M, ed. *The Cognitive Neurosciences*. 3rd ed. Cambridge, MA: MIT Press; 2004, 853–864.
84. Paivio A. *Mental Representations: A Dual-Coding Approach*. New York: Oxford University Press; 1987.
85. Marshall J, Chiat S, Robson J, Pring T. Calling a salad a federation: an investigation of semantic jargon. Part 2–verbs. *J Neurolinguistic* 1996, 9:251–260.
86. Bedny M, Caramazza A, Grossman E, Pascual-Leone A, Saxe R. Concepts are more than percepts: the case of action verbs. *J Neurosci* 2008, 28:11347–11353.
87. Simmons WK, Ramjee V, Beauchamp MS, McRae K, Martin A, Barsalou LW. A common neural substrate for perceiving and knowing about color. *Neuropsychologia* 2007, 45:2802–2810.
88. Willems RM, Toni I, Hagoort P, Casasanto D. Neural dissociations between action verb understanding and motor imagery. *J Cogn Neurosci* 2009, 22:2387–2400.
89. Boulenger V, Roy A, Paulignan Y, Deprez V, Jeannerod M, Nazir T. Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *J Cogn Neurosci* 2006, 18: 1607–1615.
90. Hoenig K, Sim EJ, Bochev V, Hernberger B, Kiefer M. Conceptual flexibility in the human brain: dynamic recruitment of semantic maps from visual, motor, and motion-related areas. *J Cogn Neurosci* 2008, 20:1799–1814.
91. Cree GS, McRae K. Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *J Exp Psychol Gen* 2003, 132:163–201.
92. Martin A, Wiggs CL, Ungerleider LG, Haxby JV. Neural correlates of category-specific knowledge. *Nature* 1996, 379:649–652.
93. Snowden JS, Goulding PJ, Neary D. Semantic dementia: a form of circumscribed cerebral atrophy. *Behav Neurol* 1989, 2:167–182.
94. Bozeat S, Lambon-Ralph M, Patterson K, Garrard P, Hodges J. Non-verbal semantic impairment in semantic dementia. *Neuropsychologia* 2000, 38: 1207–1215.
95. Garrard P, Carrol E. Lost in semantic space: a multi-modal, non-verbal assessment of feature knowledge in semantic dementia. *Brain* 2006, 129: 1152–1163.
96. Werning M, Hinzen W, Machery E, eds. *The Oxford Handbook of Compositionality*. Oxford: Oxford University Press. In press.
97. Fodor JA. *The Language of Thought*. Cambridge, MA: Harvard University Press; 1975.
98. Fodor JA. The present status of the innateness controversy. In: Fodor JA, ed. *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT Press; 1981, 257–316.
99. Laurence S, Margolis E. Radical concept nativism. *Cognition* 2002, 86:25–55.
100. Weiskopf DA. The origins of concepts. *Philos Stud* 2008, 140:359–384.
101. Carey S. Précis of *The Origin of Concepts*. *Behav Brain Sci* 2011, 34:113–162.
102. Kaye LJ. Are most of our concepts innate? *Synthese* 1993, 95:187–217.
103. Cowie F. *What's Within*. New York: Oxford University Press; 1999.
104. Churchland P. *Neurophilosophy*. Cambridge, MA: MIT Press; 1986.

105. Putnam H. *Representation and Reality*. Cambridge, MA: MIT Press; 1988.
106. Samet J, Flanagan O. Innate representations. In: Silvers S, ed. *Rerepresentation*. New York: Kluwer Academic; 1989, 189–210.
107. Sterelny K. Fodor's nativism. *Philos Stud* 1989, 55:119–141.
108. Rey G. Concepts and stereotypes. *Cognition* 1983, 15:237–262.
109. Edwards K. What concepts do. *Synthese* 2009, 170:289–310.
110. Laurence S, Margolis E. Concepts and cognitive science. In: Margolis E, Laurence S, eds. *Concepts, Core Readings*. Cambridge, MA: MIT Press; 1999, 3–82.
111. Margolis E, Laurence S. Concepts. In: Zalta EN, ed. The Stanford Encyclopedia of Philosophy (Fall 2011 Edition). Available at: http://plato.stanford.edu/archives/fall2011/entries/concepts/.
112. Malt B. Why we should do without concepts. *Mind Lang* 2010, 25:622–633.
113. Mazzone M, Lalumera E. Concepts: stored or created? *Mind Mach* 2010, 20:47–68.
114. Slaney KL, Racine TP. On the ambiguity of concept use in psychology: Is the concept 'concept' a useful concept? *J Theor Philos Psychol* 2011, 31: 73–89.
115. Machery E. A better philosophy for a better psychology: comments on Slaney and Racine (2011). *J Theor Philos Psychol* 2011, 31:90–95.
116. Couchman JJ, Boomer J, Coutinho MVC, Smith JD. Carving nature at its joints using a knife called concepts. *Behav Brain Sci* 2010, 33:207–208.
117. Danks D. Not different kinds, just special cases. *Behav Brain Sci* 2010, 33:208–209.
118. Hayes BK, Kearney L. Defending the concept of "concepts". *Behav Brain Sci* 2010, 33:214.
119. Zaki S, Cruz J. Parsimony and the triple-system model of concepts. *Behav Brain Sci* 2010, 33:230–231.
120. Nersessian NJ. How do scientists think? Capturing the dynamics of conceptual change in science. In: Giere RN, ed. *Cognitive Models of Science*. Minneapolis: University of Minnesota Press; 1992, 5–22.

