their members *qua* members, yet each is very useful in organizing established knowledge and continuing inquiry in their respective domains. And the same holds for "concept." For example, the neo-empiricist, friendly to Machery's general take on our conceptual systems, might want to defend the substantive claim that perceptual symbols are a kind of concept. Such a claim would be substantive, to the effect that the delineated class contains some additional sorts of entities. Indeed, Machery himself wonders whether there are other kinds of concepts (p. 249). For example, he writes, "Evidence shows that people have some knowledge about ideals. What is now needed is to determine whether these bodies of knowledge qualify as concepts" (p. 249). This strikes us as a meaningful and important question, and one for which the term "concept" is obviously useful in asking.

So, independent of the question of hybrids, psychologists should keep the term. Even if Machery is right, and concepts are not a natural kind, the potential dangers here would be better addressed through reformation instead of elimination. The practical advice to take away from Machery's arguments may be, not that scientists should get rid of "concept," but that they should be more careful in understanding that this term likely fails to pick out a very tidy sort of natural kind. Doing so should allow them to steer around the sorts of dangers that Machery (2009) hypothesizes about (e.g., pp. 242–243), without sinking an otherwise fruitful vehicle of inquiry. The psychology of categorization, inference, and so on may be much messier than philosophers and psychologists have hoped. But "concept" is still likely to be a vitally important word for theorizing about that mess.

Concept talk cannot be avoided

doi:10.1017/S0140525X10000361

James A. Hampton
Department of Psychology, City University London, London, EC1V OHB, United Kingdom.
hampton@city.ac.uk
www.staff.city.ac.uk/hampton

Abstract: Distinct systems for representing concepts as prototypes, exemplars, and theories are closely integrated in the mind, and the notion of concept is required as a framework for exploring this integration. Eliminating the term "concept" from our theories will hinder rather than promote scientific progress.

While most people interested in concepts will find much to agree with in this book (*Doing without Concepts*, Machery 2009), it is the eliminativist thesis that will find most resistance. Machery provides analogical cases in psychology such as "emotion" and "memory." Emotion and memory, it is argued, may prove to be terms referring to a varied set of phenomena, without any identifiable single associated brain system. Similar cases can be found in other sciences – for example, "species" and "planet." The concept of species is problematic because there is not always a clear criterion for differentiating one species from another; instead biological laws describe the distribution of genes over populations of individuals (Mayr 1982). While problems of definition mean that "species" is not a well-defined term in biology, it would, however, be hard to imagine biological discourse without it. There are just too many general truths that need to be expressed. Similarly, astronomers ran into trouble with the designation of Pluto as a planet, given the discovery of other large orbiting bodies that had been labeled as asteroids. But the term still has a referential meaning. Science needs more loosely defined general referring expressions in addition to the carefully defined terms that figure in theories. I argue that cognitive science still needs the notion of "concept," even if it proves multifaceted and hard to define satisfactorily.

Machery's argument rests on there being three distinct forms of knowledge that are recruited by default by cognitive processes: namely, prototypes (P), exemplars (E), and theories (T). The danger of eliminating the notion of concept is that the importance of the relations between these forms of knowledge risks being underplayed. First, there is the obvious point that the P, E, and T representations (let's call them PET) of dog all refer to the same class – they are broadly co-referential (give or take some differences in categorization resulting from exceptional contexts). What makes them co-referential is the fact that they represent the same concept. Without a notion of concept, it is hard to explain why they co-refer.

More importantly, the term "concept" is needed as part of an account of the many situations in which the PET systems interact. How does one discuss concept combination, including the formation of composite prototypes, the importing of exemplar knowledge, and the coherence checking of the result through background theory, if one cannot have the integrative term "concept" to specify just what it is that is being combined. The combination occurs at the concept level, and the description of the processes involved then requires elaboration in terms of the PET systems. Similarly, in concept learning, we need an overarching notion of concept in order to describe how PET systems interact. Experiential concepts like DOG or CUP may first be learned by a child through interacting with individuals encountered in everyday life. When a variety of individuals are known, and it is necessary to learn to use the words "dog" and "cup" correctly, then prototypes may be formed, enabling generalization to other individuals, discrimination of other classes, and the accumulation of generic knowledge. As the child then develops wider knowledge, the prototype notion of DOG may be supplemented by theoretically driven concepts like mammal or species, and by essentialist ideas about the causal properties of biological kinds, or the need to defer to expert opinion about correct classification.

Far from aiding scientific advance, treating the PET systems as largely independent of each other may impede investigation of the important ways in which information is transferred between them. It can also be argued that the three systems are not as easily distinguished as Machery would require. Consider prototypes and exemplars. Machery agrees that much of the research and debate concerning prototypes and exemplars has been directed at a very restricted form of behavior, namely, learning to classify simple geometrical shapes in a laboratory setting where the categories to be learned are not easily distinguished without extensive training. Even in this arcane area of psychology, there is considerable evidence that under different conditions people will either learn individual exemplars or will abstract prototypes (Smith & Minda 1998). If we move to the more "conceptual" domain of natural language terms, then the question of prototype versus exemplar models hardly arises. For example, Storms et al. (2000) have investigated whether typicality in superordinate categories like FISH, FRUIT, or FURNITURE is best predicted by similarity to the category prototype or by similarity to "exemplars." But in this case the exemplars are simply prototypes defined at a more specific level (e.g., CHAIR and TABLE). So the question is not which of two distinct systems is driving the behavior, but rather which level of abstraction is involved within a single representational system. Some concepts do have genuine exemplars – the concept of "Beethoven Symphony" to a musician will be heavily dependent on knowledge of the nine exemplars. But there will be a close link between knowledge of the exemplars and generalized knowledge about the typical structure and expressive vocabulary found in the works.

Likewise, there has been a rapprochement between prototype and theory-based elements of concepts. In discussing the notion of prototype (Hampton 1998), I have proposed that the distinguishing feature of prototype representations is that they represent the center of a class and not its boundary. It is this fact that gives rise to category vagueness, typicality gradients, the lack of explicit definitions, and the preponderance of generic (rather than necessary) features in people's accounts of the content of the concept. The notion of prototype as a form of schema is therefore free to be supplemented by causal connections within the representation resulting in a structured frame representation (Barsalou & Hale 1993). Mutability and centrality of properties, modal judgments of necessity, and dissociations between similarity-based typicality and theory-based categorization can all be accommodated within this single representational system.

In short, it is too soon to be counseling despair about integrating prototype, exemplar, and theory-based representations into a coherent account of the concept of concept.

Eliminating the "concept" concept

doi:10.1017/S0140525X10000567

Stevan Harnad
Institut des Sciences Cognitives, Université du Québec à Montréal, Montreal, Canada, H3C 3P8; School of Electronics and Computer Science, University of Southampton, SO17 1BJ Southampton, United Kingdom.
harnad@ugam.ca
http://users.ecs.soton.ac.uk/harnad

Abstract: Machery suggests that the concept of "concept" is too heterogeneous to serve as a "natural kind" for scientific explanation, so cognitive science should do without concepts. I second the suggestion and propose substituting, in place of concepts, inborn and acquired sensorimotor category-detectors and category-names combined into propositions that define and describe further categories.

Whatever a "concept" is, we have at least one for every thing we can recognize, act on, name, or describe, including not only the things denoted by all the dictionary words we understand, but also everything we know what to *do* with (Harnad 2007), even if we don't know its name or it has none – perhaps because, like "things that are bigger than a breadbox," no one has ever bothered to name it.

"Things" can be individual objects (nonliving or living), kinds, events, actions, properties, or states. We have "concepts" of countless such things, and having the concept simply means being able to *do* something with respect to those things, an action that has a right and wrong about it – anything from approaching/avoiding the thing, to interacting with or manipulating it in some way, identifying it (correctly) by name, saying true things about it, imagining it, and thinking and reasoning about it.

In *Doing without Concepts*, Machery (2009) suggests that although there is no "natural kind" corresponding to the intersection of prototypes, examples, theories, and sensorimotor representations, each may still turn out to be a legitimate natural kind of its own. I will sketch an alternative that scraps both the use and the mention of "concept" altogether.

Consider concept's twin, "percept." If a concept is, roughly, an "idea," then a "percept" is an "image." Should we ban talk of percepts, too? Pylyshyn (1973) suggested banning talk of "images" – as unobservable, unmeasurable, homuncular, and, most important, nonexplanatory – to be replaced by propositions, and, eventually, computations, which are genuinely explanatory, in that they can *generate* the capacity that the images or "percepts" had been meant to explain (Harnad 2006).

With findings on mental rotation (Shepard & Cooper 1982), however, "percept" has made a comeback, in the form of internal analog structures and processes that have some of the properties of images but can do the internal generative work, with no homunculus, sometimes more efficiently than computation. (Digital computation can always approximate analog dynamics as closely as we like: A picture is always worth more than 1,000 words, but 10,000 words come closer. It cannot, however, be words all the way down; Harnad 1990.)

Apart from their sensory shapes, objects have sensorimotor "affordances": things that objects are amenable to having done with them (by our bodies, and their shapes). A chair (but not a pyramid or a pincushion) affords sittability-upon; a doornail, but not a doormat, affords grasping and turning. But is an affordance-detector a "representation"?

We need to be able to recognize birds, for example, before we can start doing anything with them, including talking and thinking about them. No machine vision program could perform anywhere near human level using prototype-matching to recognize birds; raw example-storage would do even worse. And without those, verbal theories could not even get off the ground (because it can't be words all the way down).

So what we need first is not bird representations, but *bird-detectors.* For most of us, visuomotor contact is our first introduction to birds, but it is not "we" who pick up the affordances; we are no more aware of the tuning of our internal category detectors than subjects in mental-rotation experiments are aware of rotating their inner images. Internal mechanisms do this "neo-empirical" work for us (Barsalou 1999; Glenberg & Robertson 2000). The work of cognitive science is to discover those mechanisms. That done, it no longer matters whether we call them concepts, ideas, notions, representations, beliefs, or meanings.

Cognitive science has not yet done this job, though Turing (1950) set the agenda long ago: Scale up to a model capable of *doing* everything we can do (Harnad 2008). The first hurdle is sensorimotor category detection: the mechanism for learning categories from sensorimotor interactions with the world, guided by error-correcting feedback. We share this capability with most other species: learning to detect and act upon sensorimotor affordances. *To categorize is to do the right thing with the right kind of thing* (Harnad 2005).

Some categories are innate: We recognize and know what to do with them because natural selection already did the "learning" by genetically pretuning our ancestors' brains. But most categories we have to learn within our lifetimes, including everything named and described in our dictionaries plus many things, actions, events, properties, and states we never bother to name: We learn to do the right thing with them, and perhaps describe them, on the fly. How did we *get* those names and descriptions? Our species is the only one that has them.

According to our account so far, we only have the categories for which we have learned through direct experience what to do with their members. One of the most adaptive things our species alone does with many of our categories is to *name* them. For, with language evolved our capacity to produce and understand strings of category names that encode truth-valued propositions, *predicating* something about something. This allowed us to acquire new categories not only by sensorimotor *induction*, but also by verbal *instruction.* For once we have a set of categories "grounded" directly in our sensorimotor capacity to detect their members and nonmembers, we can also assign each category an agreed, arbitrary name (Harnad 1990), and then we can define and describe new categories, conveying them to those who do not yet have them, by combining and recombining the names of our already grounded categories (Cangelosi & Harnad 2001) in propositions. Then and only then does the "theory-theory" come in, for verbal definitions and descriptions are higher-order category-detectors, too, as long as all their component terms are grounded (Blondin-Massé et al. 2008). Here we are right to call them "representations," for they are descriptions of categories, and can be given to and received from others without every individual's having to learn the categories directly from experience – as long as the category-names used in those descriptions are ultimately grounded in direct sensorimotor categories.

There is much ongoing research on the mechanisms of sensorimotor category learning in computers, neural nets, robots, and the brain, as well as on the origins and mechanisms of natural language processing. It is nowhere near Turing-scale, but this sketch rearranges the cognitive landscape a bit, to preview how we can, as Machery suggests, *do* "without concepts": What takes their place is innate and mostly learned sensorimotor category-detectors (for which the learning mechanisms are still not known, but