Commentary on Machery [Book Review, BBS 33(2/3), 2010]

Eliminating the "concept" concept

Stevan Harnad

Canada Research Chair in Cognitive Sciences,
Université du Québec à Montréal, Montreal, Canada H3C 3P8
&
School of Electronics and Computer Science, University of Southampton SO17 1BJ
harnad@uqam.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 doorknob, but not a doornail, 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 neither prototypes nor exemplars are likely to play much of a role in them), progressively supplemented by verbal category representations composed of grounded category names describing further categories through propositions. The real challenge is getting this to work, Turing-scale. Alongside that momentous and substantive task, which of the landmarks we elect to dub "concepts" or "ideas" seems pretty much a matter of taste. [A fuller version of this commentary, entitled "Concepts: The Very Idea" is available at http://eprints.ecs.soton.ac.uk/18029]

Barsalou, L. W. (1999) Perceptual symbol systems. *Behavioral and Brain Sciences* 22(4):577–660. [SH]

Blondin-Massé, A., Chicoisne, G., Gargouri, Y., Harnad, S., Picard, O. & Marcotte, O. (2008) How is meaning grounded in dictionary definitions? Paper presented at the TextGraphs-3 Workshop – 22nd International Conference on Computational Linguistics, 18 August 2008. [SH]

Cangelosi, A. & Harnad, S. (2001) The adaptive advantage of symbolic theft over sensorimotor toil: Grounding language in perceptual categories. *Evolution of Communication* 4(1):117–14. [SH]

Glenberg, A. M. & Robertson, D. A. (2000) Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language* 43:379–401. [SH]

Harnad, S. (1990) The symbol grounding problem. *Physica D* 42:335–46. [SH]

Harnad, S. (2005) To cognize is to categorize: Cognition is categorization. In: *Handbook of categorization,* ed. C. Lefebvre & H. Cohen, 20-42. Elsevier. [SH]

Harnad, S. (2006) Cohabitation: Computation at 70, cognition at 20. In: *Essays in Honour of Zenon Pylyshyn,* ed. D. Dedrick, 245-257. MIT Press. [SH]

Harnad, S. (2007) From knowing how to knowing that: Acquiring categories by word of mouth. Paper presented at the Kaziemierz Naturalized Epistemology Workshop (KNEW), Kaziemierz, Poland, September 2, 2007. [SH]

Harnad, S. (2008) The annotation game: On Turing (1950) on computing, machinery and intelligence. In: *Parsing the Turing test: Philosophical and methodological issues in the quest for the thinking computer,* ed. R. Epstein & G. Peters, pp. 23–66. Springer. [SH]

Machery, E. (2009) *Doing without concepts.* Oxford University Press. [SH]

Pylyshyn, Z. W. (1973) What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin* 80:1–24. [SH]

Shepard, R. N. & Cooper, L. A. (1982) *Mental images and their transformations.* MIT Press/Bradford. [SH]

Turing, A. M. (1950) Computing machinery and intelligence. *Mind* 49: 433–460. [SH]