Concepts: The Very Idea

(Review of Machery (2009) Doing Without Concepts. OUP)

Chaire de recherche du Canada
Institut des sciences cognitives
Université du Québec à Montréal
Montréal, Québec
Canada H3C 3P8
http://www.crsc.uqam.ca/
&
Department of Electronics and Computer Science
University of Southampton
Highfield, Southampton
SO17 1BJ UNITED KINGDOM
http://www.ecs.soton.ac.uk/~harnad/

ABSTRACT: *Machery (2009) 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.*

Concepts, ideas, notions. Our vocabulary is rife with instances of the word "concept." To a first approximation, it means "idea". The trouble is that apart from the fact that ideas occur in people's minds, and hence in their heads, "idea" is far too vague a notion to be useful beyond casual everyday folk chatter. In fact, "notion" (and a lot of other similarly vague synonyms) can also be substituted for every mention of the word "concept." I challenge anyone to find a substantive difference between saying "I have an idea -- or a concept -- or a notion" of *apple* or of *an apple* or of *what an apple is*, or of *what the word "apple" means*, or even of *what the word "apple" refers to*. In fact, the notion of "concept" is very closely tied to the notion of word meaning. If I know what the word "X" means (regardless of whether X is a content word or a function word) then chances are I have a concept of X. (But what is a "concept"?)

(Don't even begin to ask about concept's adjectival form, "conceptual," let alone its adverb, "conceptually," because those become virtually syonymous with "mental," "intellectual," or just plain "cognitive," and those are all notoriously nonexplanatory terms, seeking, rather than providing explanation.)

What do you have if you have an idea? So a concept is more or less an "idea." What does someone have, if he has a concept or an idea of an apple? It usually means he can *recognize* an apple if he sees one. He can *imagine* what it looks like. And he can tell you both what an apple looks like (red and round) and what it is (the fruit of an apple tree); in other words, he can *describe or define* it (usually to a close enough approximation so that – unless you cannot speak the language or you are from another planet -- you too will be able to recognize, imagine, describe and define an apple once he's done telling and perhaps also showing you).

Doing and telling. So we've established that, whatever a concept is, we have at least one of them for every thing and kind of thing that we are able to recognize, name or describe, including not only the things denoted by all the dictionary words we understand, but also every thing and kind of thing that we know what to *do* with – behavioral "know-how" (Harnad 2007) -- even if we don't know its name or it hasn't got one – perhaps because, like "things that are bigger than a breadbox," no one has ever bothered to name it.

Objects, events, actions, properties or states. "Things," by the way (another hopelessly vague word) can be individual objects, nonliving and living, kinds of objects, 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. ("Thinking," along with "things" and "ideas," being suitable only for folklore, on account of their extreme vagueness and lack of substance.)

"Natural kinds." You may be sensing now that I am about to become even more dismissive of the concept of "concept" than Edouard Machery has been, and you are right! Machery's case against concepts is that they fail to constitute a "natural kind." (For those of you who lack a concept of "natural kinds," they are supposed to be the kinds of things that scientists find useful in explaining nature. An electron would be a natural kind, whereas "things that are bigger than a breadbox" would not be.) Machery examines the various theories of both psychologists and philosophers about what concepts are and concludes that all of them together do not have enough in common to warrant continuing to use the word "concept" for these diverse entities.

Prototypes, examples, theories, sensorimotor representations. Psychologists have concepts that are "prototypes" of the things they are concepts of; they also have concepts that are examples of them; concepts that are theories of them; and even concepts that are sensorimotor "representations" of them. Some philosophers have concepts of concepts that have been allegedly invalidated by Wittgenstein, not to mention Eleanor Rosch (Rosch & Lloyd 1978), for having claimed that whether a thing falls under a concept depends on its fulfilling certain necessary and sufficient conditions. There are allegedly few things that meet this exacting criterion, so this

so-called "classical" concept of concept has been declared, by some, to be incorrect, replaced instead by the various prevailing views (prototypes, examples, theories, sensorimotor representations, etc.). There is also some difference of opinion on whether we are allowed to have one concept of a thing, or many.

Uses and mentions of "concept." For all these reasons and more, Machery has concluded that there is no useful invariant underlying all these uses of the word and concept of concept, so it ought to be dropped -- though of course Machery continues to use it, and not just mention it, throughout his book, just as I have been doing here. It would be an illuminating exercise – feasible with some digital processing – to replace with "X" in Machery's book all uses of the word "concept" (as opposed to mentions of it, in quotes) in order to see how many of the book's assertions would still be intelligible. But never mind; we are entitled to continue using "prescientific" folknomics to keep going while we wait for the science.

Multiple representations. What about the science, then? Machery is far more impressed than I am with the progress in the cognitive science of X (where X used to be concepts). He is impressed enough to allow that whereas there is no "natural kind" corresponding to the intersection of prototypes, examples, theories and sensorimotor representations, each of these may still turn out to be a legitimate natural kind of its own. Machery welcomes that also because he thinks it's important that we may have several different what-we-used-to-call-concepts of the same thing, and now that we don't have concepts any more, we can speak of having prototypes of things, examples of things, theories of things and sensorimotor representations of things, and that way empirical experiment and theory can progress without being constrained to conceive (there's that word again!) and pursue all these approaches as if they were all part of the same grand unified theory of concepts.

Percepts? I will now sketch a rival approach, one that, in accord with Machery's suggestion, scraps both the use and the mention of "concept" altogether. The word "concept" is already banned in my lab: If anyone utters it, they must immediately replace it by what they actually mean by it, and invariably what they mean is the unknown cerebral wherewithal that generates the capacity to *do* something or other; and that something is usually to categorize something in a certain way. But I will return to this. First, let us consider concept's twin, "percept."

Sensorimotor images. What is a percept? If a concept is, roughly, an "idea," then a "percept" is probably, roughly, an "image," although the "image" can be in any sensory or motor modality (not just visual), and it includes not only what we imagine, but what we see, hear, taste, smell and feel when we are actually perceiving something. In other words, a "percept" is either an actual, ongoing, or a remembered or imagined sensorimotor experience.

Propositions, computations, explanations and homunculi. So if there is now talk of banning "concept" from our professional vocabulary, what about its perceptual homologue and twin, "percept"? Well there has of course been a good deal of talk of banning talk of "images" too, notably by Zenon Pylyshyn (1973), on the grounds that

images are unobservable, unmeasurable, homuncular, and, most important, nonexplanatory. They just reduplicate the problem of explaining perception with an internal copy of the external thing we perceive, to be in turn perceived by an internal homunculus, without ever explaining how it's done. Talk of images was meant to be replaced by talk about propositions, and, eventually, computations. These, according to the critics, were actually explanatory, in the sense that they could generate the capacity that the images or "percepts" had been meant to explain, by actually generating sensorimotor performance, producing, given our sensory inputs and tasks, the outputs that we ourselves produced, facing the same inputs and tasks (Harnad 2006).

Mental Rotation. Now there is a lot to be said for this change in methodology, in favor of "concepts" that are actually predictive, generative, and explanatory. But thanks to Roger Shepard's findings on mental rotation (Shepard & Cooper 1982), the concept of "percept" has also been making a comeback, in the form of internal analog structures and processes – hence not necessarily propositional or computational ones – that do have some of the properties we had attributed to mental images, and yet they actually do the internal generative work, without the need of a homunculus, and often more efficiently than digital computation would have done. (It is true that digital computation can always approximate analog dynamics as closely as we like: Yes, a picture is always worth more than a thousand words, but ten thousand comes closer. It cannot, however, be words all the way down; Harnad 1990.)

"Neo-Empiricism." This comeback on the part of percepts in the form of internal analog processes is also present in the fourth of the current approaches to concepts discussed by Machery, namely, the embodied sensorimotor representations of Barsalou (1999), Glenberg (Glenberg & Robertson 2000) and others that Machery calls "neo-empiricist."

"Neo-Behaviorism." This brings me to another ubiquitous term of art in cognitive science -- one that few have yet considered proposing to banish (although the behaviorists tried to do so in advance, until it was they who were banished): "representations." This word too is banished from my lab. It too must be replaced by whatever structure or process is being claimed to be embodied, encoded or otherwise implemented in the brain, and in the service of what input and task. This is not neo-behavioriorism. The behaviorists banned, in one swoop, all talk of what was not empirically observable in the form of input and output. That included not just the unobservables one might, from an armchair, introspect as going on in one's mind, but also the unobservables one might hypothesize, on the basis of evidence in the lab, to be going on in one's brain, or going on inside any system capable of doing the sorts of things we are capable of doing.

Mental representations. Now what's wrong with using "mental representations" as short for "whatever internal structures and processes generate whatever cognitive capacity we are trying to explain"? Nothing, except, again, the homuncularity inherent in the notion of "representing." For, just as it becomes

apparent in the case of Shepardian "mental rotation" (which is not really *mental* rotation at all, but *cerebral* rotation, for it occurs much too fast for us to be aware of it or to do it deliberately), that the rotation is not a "representation" of anything to anyone, least of all to some inner homunculus, but simply an internal operation on input that generates the right output, so in the case of the various mental representations posited by the four approaches to concepts reviewed by Machery – prototypes, examples, theories, sensorimotor representations – one wonders what work the notion of "representing" is actually doing.

Sensorimotor activity. Consider sensorimotor representations, the ones that wear the term "representation" on their sleeves: A typical experiment in this line of research would be to show that an observer's recognition of a picture of something, or the name of something, is facilitated if you get the observer to perform the kind of action that he would normally perform on the thing depicted or named, and his recognition is impaired if you get him to perform an action that is incompatible with the action that he would normally perform on the thing depicted or named: pulling a wagon versus pushing a shopping-cart. This does suggest that in recognizing a wagon or the name of a wagon, some motor activity may be involved, but is that a "representation" of a wagon, as if your brain were pulling on an inner wagon in order to determine whether the picture or word it is viewing is the picture or the name of a wagon? I'm sure sensorimotor activity of some sort is involved in such tasks, but "representation" might be not only a misnomer, but misleading, as to what the sensorimotor activity really is.

Detecting sensorimotor "affordances." Again, the work of Shepard (1984) – who became increasingly Gibsonian across the years – gives us a clue: Objects don't just have sensory shapes; they also have sensorimotor "affordances": Affordances are 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 doormat, affords grasping and turning. So it would be foolish of the brain not to make use of the cues to these sensorimotor affordances in recognizing objects. But is an affordance-detector a "representation"? (Nor does "simulation" quite capture what is going on here; nor does Barsalou's (1999) "perceptual symbolization."_

Prototype-matching. Perhaps you think a prototype is a better example of a mental representation? In principle, it would be, if it turned out that most or many things are actually recognized on the basis of prototype-matching, whereby the shape of an external object is compared to the shape of stored inner prototype shapes, and the object is recognized as being of the kind whose internal prototype it matches most closely.

Category Recognition. Now, as with Shepard's mental rotation experiments, no observer is ever aware of matching his sensory input to a series of internal prototypes. When I recognize a bird as a bird (or the name of a bird as the name of a bird), I am not aware of trying to match its shape with that of the stored shapes of birds or fish or chairs. All I'm aware of is that when I think of birds (especially when

the thinking is triggered by someone's saying "bird") I tend to have a certain typical bird in mind, closer to a sparrow or robin than an ostrich or a penguin. No doubt I recognize typical birds more quickly (but one need not hypothesize prototype representations for that: the rival example-hypothesis could explain that even better just in terms of frequency of prior exposure). No doubt I would rate typical birds as more typical (but that does not make the atypical ones less truly and fully birds!). And no doubt I am unable to give a definitive definition or an exhaustive description of what a bird is. Yet I can reliably recognize as a bird all birds I see, whether typical or atypical. How do I do it? The prototype theory excels for the typical birds. The example theory does better for the less typical birds. And the "theory theory" can resolve even the most atypical cases if the theory is elaborate enough (e.g., 10,000 words!). Which is the "representation" of "bird"?

"Heterogeneous" representations. Machery might reply "all of them," because there are multiple, heterogeneous concepts of "bird" – or, if we obey, and abandon the word "concept," we have multiple "representations" of the ex-concept "bird." But do we even have that? Can we really put together those heterogeneous "representations" into something that can do a serviceable chunk of what we can do vis-a-vis birds, and vis-a-vis discourse about birds (Harnad 1987)? Something that can recognize, name, interact with, manipulate, define, describe, discuss, and understand discourse about birds? Prototypes fail completely. No machine vision program could perform at anywhere near human level if it used prototype-matching to recognize birds; raw rote example-storage would do even worse. And without those, the verbal theories could not even get off the ground (because it can't be words all the way down).

Category-detectors. Add sensorimotor representations? A good idea. In fact, a better idea would be to forget about prototypes and examples altogether, hold theories in abeyance until there's something for their words to hook onto, and *begin* with sensorimotor representations – but, as already noted, it's not "representations" we're looking for. In the first instance, we're looking for *recognition*: We need to be able to recognize birds before we can start doing anything with them, including talking and thinking about them. So what we are looking for is not bird representations, but *bird-detectors*. And "neo-empiricist" sensorimotor bird-detectors are a good place to start, because for most of us, visuomotor contact is our first and critical introduction to birds. It may be from walking tours of the neighborhood, with mum pointing out and naming the various things, living and nonliving, bird and nonbird that we encounter, or it may be the same thing with mum and a picture book. With birds, visual examples, pointing, naming, and error-corrective feedback may be enough to get us launched. With chairs, wagons and shopping-carts, the sensorimotor interaction will be more interactive than just visual and verbal, so that we can pick up the sensorimotor affordances.

The internal machinery of cognition. Except that it is not "we" who pick up the affordances; for the most part, we are no more aware of the tuning of our internal category detectors than the subject in the mental rotation experiment is aware of

the rotation of his inner analogs. We have internal structures and processes that do this neo-empirical work for us, and the work of cognitive science is to discover what those internal structures and processes are. That done, it no longer matters a whit whether we choose to call them concepts, ideas, notions, representations, beliefs, meanings, or what have you. They are the internal machinery of cognition, whatever it turns out to be.

The Turing Test. Cognitive science has not yet done this job, although Alan Turing already set the agenda over a half century ago (Turing 1950): The goal is to scale up to a model that is capable of doing everything that we are capable of doing -- otherwise known as passing the Turing Test (Harnad 2008).

Sensorimotor category detection. So let us fantasize a bit. The first hurdle is sensorimotor category detection. That means cognitive science needs to discover the structures and processes that are capable of learning categories on the basis of sensorimotor interactions with the world, under the guidance of error-correcting feedback from the consequences of the sensorimotor interactions, adaptive and maladaptive. This is a form of category learning that we share with most other species that are capable of learning at all. One can think of it as learning to detect and act upon sensorimotor affordances. The act itself can be described as categorization in a very general sense: *To categorize is to do the right thing with the right kind of thing* (Harnad 2005).

Category learning. Some categories – or rather, their pretuned category-detectors -- are innate: We recognize them, and know what to do with them, without having had to learn anything. Natural selection did the "learning" by genetically pretuning our ancestors' brains. But for most categories we have to learn them in our lifetimes: That includes just about all of the items catalogued in our dictionaries and encyclopedias -- not to mention the many kinds of things, actions, events, properties and states that we never bother to immortalize with a name: We simply learn to do the right thing with them, and perhaps describe them on the fly.

Categories by induction. But 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 those categories with which we have learned to do the right thing on the basis of direct trial-and-error experience (and imitation), with the help of our tunable category-detectors: whether to approach or avoid them, eat or escape from them, mate with them, compete with them, nurture them, sit on them, or do any of the other sensorimotor things with them that they afford doing and that are adaptive for us.

Categories by instruction. For our species alone, however, one of the most adaptive things we do with a great 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 of something. This allowed us to acquire new categories not only by sensorimotor induction, but by verbal instruction. Once we have -- via our tunable category detectors -- 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). 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). And here we are right to call them "representations," for they are descriptions of categories, and can be given to and received from others without everyone's having to learn the categories for himself from direct trial and error sensorimotor experience – as long as the category-names used in those descriptions are ultimately grounded in direct sensorimotor categories.

Doing without concepts. I have not, of course, explained either how learned category-detectors work, nor how language (the capacity to produce and understand any and every proposition) works. There is a good deal of ongoing research on the mechanisms of sensorimotor category learning in computers, neural nets, robots and the brain, as well as a good deal of research on the origins and mechanisms of natural language processing. It's nowhere near Turing-scale, but I just wanted to rearrange the cognitive landscape a little, 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 it is pretty sure that neither prototypes nor exemplars will play a very decisive role in them), progressively supplemented by verbal category representations composed out 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.

REFERENCES

Barsalou, L.W. (1999) Perceptual symbol systems. *Behavioral and Brain Sciences* 22 (4) 577-660

Blondin-Massé, A, Chicoisne, G, Gargouri, Y, Harnad, S, Picard, O & Marcotte, O(2008). How Is Meaning Grounded in Dictionary Definitions? In *TextGraphs-3 Workshop - 22nd International Conference on Computational Linguistics*

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

Glenberg, Arthur M. and David A. Robertson (2000) Symbol Grounding and Meaning: A Comparison of High-Dimensional and Embodied Theories of Meaning. *Journal of Memory and Language* 43: 379–401

Harnad, S. (1987) Category Induction and Representation, Chapter 18 of: Harnad, S. (ed.) (1987) *Categorical Perception: The Groundwork of Cognition* . New York: Cambridge University Press.

Harnad, S. (1990) The Symbol Grounding Problem. *Physica D* 42: 335-346.

Harnad, S. (2005) To Cognize is to Categorize: Cognition is Categorization, in Lefebvre, C. and Cohen, H., Eds. *Handbook of Categorization*. Elsevier.

Harnad, S. (2006) Cohabitation: Computation at 70, Cognition at 20, in Dedrick, D., Eds. *Essays in Honour of Zenon Pylyshyn*

Harnad, S. (2007) From Knowing How To Knowing That: Acquiring Categories By Word of Mouth. Presented at *Kaziemierz Naturalized Epistemology Workshop* (KNEW), Kaziemierz, Poland, 2 September 2007.

Harnad, S. (2008) The Annotation Game: On Turing (1950) on Computing, Machinery and Intelligence. In: Epstein, Robert & Peters, Grace (Eds.) *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Springer

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

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

Rosch, E. & Lloyd, B. B. (1978) *Cognition and categorization*. Hillsdale NJ: Erlbaum Associates

Shepard, RN (1984) Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review* 91(4): 417-447.

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

Turing, A.M. (1950) Computing Machinery and Intelligence. *Mind* 49: 433-460