A Refutation of the Circularity Threat in Concept Acquisition

José V. Hernández-Conde[1][0000-0002-8502-6570]

[1] University of Valladolid, Pza. Campus Universitario s/n, 47011 Valladolid, Spain
jhercon@uva.es

Keywords: Concept Acquisition, Circularity, Empiricism, Nativism.

One of the main problems of concept empiricism, a theory according to which all concepts are learned –not innate–, is to explain the acquisition of primitive concepts without resorting to preexisting innate elements. According to nativism, an innate representational repertoire (on which later learning is based) is one of the key elements in every theory of concept acquisition. However, this is not a choice for a consistent empiricist, because it compels to the acceptance that concept learning consists in the production of complex concepts from a set of innate elements. Therefore, it is required a response from empiricism to the reasons provided by nativists [1] against the empiricist claim that primitive concepts may be acquired.

Fodor claims that empiricism is wrong because concepts cannot result from assembling their definitional constituents:

- [P_1] Concepts –either primitive or complex– cannot be acquired by hypothesis formation and confirmation (HF).
- [P_2] There are no other learning methods.
- [C_1] Concepts cannot be learned.

This argument was later expressed by Carey [2] in a clearer way:

- [P_I] All learning mechanisms reduce to hypothesis formation and testing (HF).
- [P_{II}] Hypotheses that play a role in the learning of new concepts must be formulated in terms of the available concepts, and the principle of compositionality.
- [P_{III}] Primitive concepts are not formulable –as definitions or probabilistically–, in terms of other concepts.
- [C_I] Therefore, primitive concepts cannot be learned.

Nonetheless, Carey's formulation of Fodor's argument can be expressed independently of the type of learning (or, in other words, without assuming that all cases of learning can be reduced to HF):

- [P_I'] Concept acquisition is the result of learning processes.
- [P_{II}'] The perceptual –or conceptual– constituents of concepts must be available before the beginning of the learning process, in order to be used as an input of its computational apparatus.
- [P_{III}'] Consequently, those constituents should have been acquired at an earlier time.

[P_{IV}'] A model like this cannot explain the acquisition of those constituents without falling into circularity. (The reason is that such constituents should result from a learning process in whose beginning another set of more basic elements has to be available, and so on and so forth.)

[C_I'] Therefore, primitive concepts cannot be learned.

The above reformulation of the argument against empiricism is more general than the ones of Fodor and Carey, because it does not depend on the kind of learning. The issue is that there is a circularity threat if both concepts and their constitutive elements are acquired by means of the very same kind of cognitive process, because the learning process ends up with an infinite regress when trying to explain how the most basic constituents of concepts are acquired.

Once reformulated that way, it is clear that the origin of the circularity problem is premise [P_{II}'] or, more specifically, the underlying *precedence assumption*:

(PA) The perceptual or conceptual constituents of concepts must be available as an input of the cognitive processes that lead to the acquisition of those concepts.

This is a common assumption of both empiricists and moderate nativists. Additionally, that is the very precise idea which underlies Fodor's reasons against concept empiricism, when he argues that if a primitive concept C_P is available to a subject S for hypothesis formation, then S already has that concept, because in the absence of C_P the subject cannot be able to formulate the tested hypothesis, nor even to recognize the experiences associated with C_P.

Under the precedence assumption, and in order to avoid the circularity threat, an empiricist has to accept the innateness of primitive concepts, what explains the common shift of empiricists towards moderate nativism.

Should the empiricist give up and accept that primitive concepts cannot be learned? On my view all is not lost for the empiricist, because he is not under the obligation to accept the precedence assumption. Thus, the thesis I will argue for is that, for a model (of concept acquisition) to be able to function, it is not necessary that its constitutive elements are available from the very beginning. By contrast, my point will be that it is enough that those most basic constituents are ready at the end of the learning process, and not from its start.

Lastly, my proposal is a model where the constitutive elements of a concept result from the very same cognitive process by virtue of which that concept is acquired. This kind of model would consist in a three-step iterative learning process:

1. *Dimensionality reduction*: the information received by the sensorium is so huge that it cannot be persistently stored, nor directly analyzed. Thence, a reduction in the number of dimensions is required, that describes most of the original data variability in terms of a reduced number of factors –removing as much redundant information as possible. This first step will allow identifying the most basic constituents of concepts.

   Besides, dimensional reduction capacities are quite plausible, which is supported by the increasing neurophysiological and neuropsychological evidence of sensory systems where dimensionality reduction is hard-wired into the brain, which

vary from the low-dimensional representations of the visual inputs [3], to the compression of dimensions carried out in the cortical processing of sensory information [4], and in sensory and motor processing of postural control [5].

2. *Pattern recognition*: regularities in the reduced data input are singled out in order to pin-point similar future stimulus in comparable circumstances. Those regularities are acquired in an unsupervised way, and can be identified with the concepts of our mental system. In this case, the pattern identification phase could take the form of a cluster analysis carried out over the output of the dimensionality reduction stage [6], [7].
   Furthermore, the concepts resulting from a cluster analysis could perfectly do the explanatory work expected of prototypes, bundles of exemplars, and proxytpes.
3. *Evaluation and readjustment*: since nothing guarantees that the obtained factors are the best ones for the subsequent pattern recognition, nor that the identified regularities are the most explanatory ones, an iterative process is required. In this stage, the dimensions and patterns –or, alternatively, the features and concepts– are evaluated in terms of their causal explanatory power, and the iterative learning process is rearranged accordingly.
   This kind of feedback is in line with the *functionality principle* [8], according to which the learned categorizations (i.e., the regularities identified in the pattern recognition stage) should have an influence on the set of extracted factors; with the substantial evidence of top-down effects [9], where lower brain regions are modulated by higher regions; and with other examples of the influence of high-level cognition over lower level conceptual / perceptual processes [10].

In a model like this the external raw data are not the most basic constituents of the conceptual system –the external data cannot even be persistently stored. Instead, the primitive elements of concepts are the factors resulting from the dimensional reduction. And, since the constitutive elements of concepts are an output –not an input– of the learning process, this model is able to explain the acquisition of concepts in a non-circular way, without resorting to a preexisting innate repertoire.

References

1. Fodor, J. A.: LOT 2: The language of thought revisited. Oxford University Press, Oxford (2008).
2. Carey, S.: The origin of concepts, Oxford University Press, Oxford (2009).
3. Pang, R., Lansdell, B. J., Fairhall, A. L.: Dimensionality reduction in neuroscience. Current Biology 26(14), R641–R666 (2016).
4. Bar-Gad, I., Morris, G., Bergman, H.: Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. Progress in Neurobiology 71(6), 439–473 (2003).
5. Ting, L. H.: Dimensional reduction in sensorimotor systems: A framework for understanding muscle coordination of posture. Progress in Brain Research 165, 299–321 (2007).

6. Blashfield, R. K., Aldenderfer, M. S.: The methods and problems of cluster analysis. In: Nesselroade, J. R., Cattell, R. B. (eds.) Handbook of multivariate experimental psychology. 2nd ed., pp. 447–473. Plenum Press, New York (1988).
7. Grabmeier, J., Rudolph, A.: Techniques of cluster algorithms in data mining. Data Mining and Knowledge Discovery 6(4), 303–360 (2002).
8. Schyns, P. G., Goldstone, R. L., Thibaut, J.-P.: The development of features in object concepts. Behavioral and Brain Sciences 21(1), 1–54 (1998).
9. Rolls, E. T.: Top-down control of visual perception: Attention in natural vision. Perception 37(3), 333–354 (2008).
10. Goldstone, R. L., Leeuw, J. R., Landy, D. H.: Fitting perception in and to cognition. Cognition 135, 24–29 (2015).