

Concept Acquisition without a Final Criterion

Michael Strevens

Draft of February 2003

7. Concept Acquisition without a Final Criterion

The absence of a final criterion poses an interesting problem about concept acquisition. On the standard stories about concepts, a person acquires a concept, and so constructs the corresponding psychological category, by stipulating, learning, or otherwise putting in place a final criterion for the category.

On the Lockean and prototype theories, the final criterion is assembled from various observable properties, by putting the properties together to form either a definition or a prototype. Note that this operation is not in itself an inference or a judgment. By including the property of yellowness in my concept for gold, I am not inferring that gold is yellow; I am as it were making it so. Of course, the *basis* for including yellowness in the concept GOLD might be inductive; it might be, for example, the observation that certain properties, among them yellowness, tend to co-occur in nature. (Rosch often argues that prototypes are constructed to capture correlations of observed properties.) But the construction of a definition or prototype is not itself a piece of inductive or any other kind of reasoning.

The essentialist theory adds, as usual, an interesting and suggestive twist to the story. A typical acquirer of an essentialist concept does not know the form of the relevant final criterion, that is, the nature of the relevant real essence. Essentialist concepts are therefore usually acquired not by constructing a final criterion, but by constructing something to stand in for the criterion: the so-called essence placeholder. As noted above, the placeholder is not a criterion itself, but rather something that commits the inferrer to using the real essence as a final criterion so far as is possible.

An essence-based category is created when a person observes a correlation in nature. The creation of the category has two stages. First, an essence placeholder is built; then, appropriate causal laws are postulated linking the essence to the relevant observable properties.

The first of these steps is the actual construction of the concept; the second step requires that the concept already exist, since, to make a posit such as *having the tiger essence causes an organism to behave ferociously* I must already have the concept of a tiger essence and so the concept of a tiger. I obtain both concepts, of course, from the first step, the construction of the essence placeholder.

Is this step an inference, a judgment, a piece of reasoning? Some would say yes. The essentialist view is usually considered to be a variant of the “theory-theory” of concepts, on which to have a concept is to accept some kind of theory. The essence placeholder might therefore be interpreted as a fundamental postulate of an appropriate theory. In the case of tigers, for example, the placeholder might be the hypothesis that “all tigers share a common essence that is causally responsible for their characteristic observable

properties". Acquiring the concept would then be equivalent to making the theoretical postulate.

A famous argument due to Fodor (1981) shows that this view cannot be quite correct. In order to make a postulate, one must have all the concepts necessary to formulate the postulate. But the thesis in question contains the tiger concept, by assumption not yet constructed. Could it be that the postulate is instead something like "all animals that are striped, ferocious, etc. share a common essence that is causally responsible for those properties"? No, because that postulate is not a part of our theory. On the essentialist view, we hold that some animals that have the listed features do not have the tiger essence; lions made up by the Keil lab to look like tigers provide one kind of example. The essence placeholder must be created, then, and so essentialist concepts acquired and the essentialist theory put in place, by some other, non-inferential process. I leave further investigation to the essentialists; I want to now consider the question how concepts are acquired if there are no final criteria.

On the minimalist picture, there are no final criteria associated with a concept, only normal, factual beliefs involving the concept, principal among which are the causal laws of the form *something about tigers causes them to be ferocious*. To acquire a concept, then, must be to acquire some of these beliefs. (It does not matter which beliefs are acquired; for example, there is no particular set of causal laws that must be believed in order to have the concept TIGER.)

But now we face squarely a problem that essentialism only narrowly avoided above: the beliefs that must be acquired in order to come to have

the tiger concept all concern tigers. Thus, to acquire the beliefs the reasoner must first have acquired the tiger concept. The tiger concept cannot, it seems, be acquired by coming to have certain beliefs about tigers; therefore, minimalism is wrong.

If the all-important beliefs can be acquired only by learning or some other form of inference, then this objection is devastating to minimalism. It is possible, however, to imagine a belief acquisition mechanism that makes room for minimalism, by forming both the beliefs about tigers and the tiger concept in a single operation. Here is how the mechanism would work.

Suppose that a learner notices that many organisms share a certain cluster of observable properties: they are striped, ferocious, fearfully symmetrical, and so on. Call these the characteristic properties. The observation of the correlation between the characteristic properties triggers in the learner a process that might be schematically conceived of as follows. (Some additional details of the acquisition mechanism are discussed in Strevens (2001).)

1. For each of the characteristic properties P , a representation of the following form is created: *Something about being a ___ causes an organism to have P* . What is represented at this stage, then, is not so much a law as a law schema.
2. A new mental predicate is created, a mental term, that is, without any cognitive significance, like a coined word. This predicate will, by the end of the process, have become the mental term for *tiger*.
3. The new predicate is dropped into the causal law schemas. For each of the characteristic properties P , then, the user now represents a law of

the form *Something about being a tiger causes an organism to have P* .

The new predicate is not given significance before being used in the laws, then, but equally, the laws are not postulated before the new predicate is given significance. The laws are affirmed and the tiger predicate given significance at the same time, by a single process.

There is nothing especially odd about this process; however, it is important to insist that the laws are not learned, inferred, or even, strictly speaking, posited. The operation by which they are created does not fit neatly into any folk psychological category.

How plausible is this story about concept acquisition? Insofar as there is reason to prefer causal minimalism as a theory of psychological categories, it had better be plausible, as there is no alternative story, as best I can see.

Still, the reader may feel queasy at my suggestion that concept acquisition goes by way of a dedicated non-inferential mechanism triggered by the observation of correlated characteristic properties. That reader should note that every theory of concepts considered in this paper must hold that concept acquisition goes by way of a process meeting this description: in the classical theory, by the process of conceptual conjunction, in the prototype theory, by the process of prototype formation, and in the essentialist theory, by whatever process constructs essence placeholders.

References

- Fodor, J. (1981). The present status of the innateness controversy. In *Representations*, pp. 257–316. MIT Press, Cambridge, MA.
- Stevens, M. (2001). Only causation matters: Reply to Ahn et al. *Cognition* 82:71–76.