pirical process) rather than experience, this is factor N, nativism. But then what are we to make of factor C, culturalism? I think that what S&B really have in mind here is what others have called "constructionism." With factors N and E, categories are derived from the structure of the world; with factor C they are somehow "constructed" by cultural practices and conventions. It is in this light that S&B introduce the "Whorf Hypothesis" (Whorf 1956), according to which our view of reality depends on our language and culture. But the Whorf Hypothesis fell on especially hard times with color categories, and S&B unfortunately inherit those hardships in using colors as their mainstay.

There are many ways in which color categories are unrepresentative of categories in general. First, they are of low dimensionality (mainly electromagnetic wave frequency, but also intensity and saturation). Second, they have a known and heavy innate component. We are born with sensory equipment that prepares us to sort (and name) colors the way we do with incomparably higher probability than the way we sort the categories named by most of the other nouns and adjectives in our (respective) dictionaries. Nor are most of the categories named by the words in our dictionaries variants on prototypes in a continuum, as colors are.

Yes, there are variations in color vision, color experience, and color naming that can modulate color categories a little; but let's admit it: not much! Moreover, color categories are hardly decomposable. With the possible exception of chromatographers, most of us cannot replace a color's name with a description – unlike with most other categories, where descriptions work so well that we usually don't even bother to lexicalize the category with a category-name and dictionary-entry at all. Even "the color of the sea" is only a one-step description, parasitic on the fact that you know the sea's color. Compare that with all the different descriptions that you could substitute for "chair."

Why does describability matter? Because it gets much closer to what language really is, and what it is really for (Cangelosi & Harnad 2001). Language is not just a category taxonomy. We use words (category names) in combination to *describe* other categories, and to *define* other words, which makes it possible to acquire categories via *instruction* rather than merely the old, prelinguistic way, via direct experience or imitation. S&B think naming's main use is to tell you which object I have in mind, out of many we are both looking at now. (It seems that good old pointing would have been enough to solve that problem, if that had really been what language was about and for.)

But not only are color categories unrepresentative of categories in general, and the joint discrimination game unrepresentative of what language evolved and is used for, but categories do not derive merely or primarily from the passive correlational structure of objects (whether picked up via species evolution or via individual experience). It is not the object/object or input/input correlations that matter, but the *effects* of what we *do* with objects: the input/output correlations, and especially the *corrective feedback arising from their consequences*. What S&B's model misses, focusing as it does on discrimination and guessing games instead of the game of life, is that categories are acquired through feedback from *mis*categorization. We have this in a realistic mushroom foraging paradigm, but not in a hypothetical discrimination/guessing game (except if we gerrymander the game so that successful discriminating/guessing becomes the name of the game by fiat, and then that is fed back in the form of error-correcting consequences).

Yet all the right elements do seem to be there in S&B's simulations. They are simply not put together in a realistic and instructive way. The task of mind-reading in context seems premature. Every categorization in fact has *two* contexts. First, there is its *context of acquisition*, in which the category is first learned (whether evolutionarily via N or experientially via E) by trial-and-error, with corrective feedback provided by the consequences of miscategorization. The acquisition context is the series of examples of category members and nonmembers that is sampled during the learning (the "training set" in machine learning terms). Until language evolves, categories can only be learned and marked on the basis of an *instrumental* "category-name" (approaching, avoiding, manipulating, eating, mating). With language, there is the new option of marking the category with an arbitrary name, picked by (cultural) convention.

When a category has already been learned instrumentally, adding an arbitrary name is a relatively trivial further step (and nonlinguistic animals can do it too). But then comes the second sense of "context": the *context of application* (for an already acquired category) in which the learned arbitrary category-names are used for other purposes. S&B's paradigm is, in fact, just one example of the context of application (telling you which of the colors that we are both looking at I happen to have in mind), but not a very representative or instructive one. Far more informative (literally!) is a task in which it is *descriptions* that resolve the uncertainty, and the alternatives are not even present. This is not discrimination but instruction/explanation. But for that you first need real language, and not just a taxonomy of arbitrary names (Harnad 2000).

What follows from this is that a "language game" in which words and categories are jointly coined and coordinated "on the fly," as in S&B's color-naming simulations, is not a realistic model for anything that biological agents ever do or did. There is still scope for Whorfian effects, but those will come from the fact that both our respective experiential "training samples" (for all categories) and our corrective feedback (for categories about which our culture and language have a say in what's what, and hence also a hand in dictating the consequences of miscategorizing) have degrees of freedom that are not entirely fixed either by our inheritance or by the structure of the external world.

Categories are *underdetermined*, hence so are the features we use to pick them out. In machine learning theory, this is called the "credit/blame" assignment problem ("which of the many features available is responsible for my successful or unsuccessful categorization?"), which is in turn a symptom of the "frame problem" (how to anticipate all potential future contingencies from a finite training sample?) and, ultimately, the "symbol-grounding problem" (how to connect a category-name with all the things in that category, past, present, and future?) Underdetermination leaves plenty of room for Whorfian differences between agents.

A synthesis of many levels of constraints as a modern view of development

Derek Harter and Shulan Lu
Department of Psychology, Computer Science and Information Systems, Texas A&M University, Commerce, TX 75429.
Derek_Harter@tamu-commerce.edu
http://faculty.tamu-commerce.edu/dharter/
Shulan_Lu@tamu-commerce.edu

Abstract: The debate of nativisim versus empiricism is over the relative importance of evolutionary versus ontogenetic mechanisms. This is mostly seen today as a false dichotomy. The synthesis of these positions provides a modern viewpoint of grounded category formation. This combined view places equal importance on feedback between these levels in guiding development, and is more appropriately compared to culturalist positions.

Much of the debate between nativism and empiricism seems to us to echo similar debates that have been prevalent in developmental psychology and biology on the question of nature versus nurture. That is to say, how much of a role does genetic evolution play in the development of behavior in humans and animals? How much can be attributed to ontogenetic learning by the individual? Is either factor predominant and, if not, are there some areas of behavior and learning where one or the other is the main contributing factor? These debates seem, however, to have reached somewhat of a preliminary consensus, that it is neither and both

https://doi.org/10.1017/S0140525X05320085 Published online by Cambridge University Press

at the same time (see, e.g., Oyama 1985), the basic idea being that both complex adaptive systems (evolution and ontogenetic development) are at once separate, but also tightly coupled with each other in mutual feedback relations. The complex and constant feedback between these levels are what define and shape the successful behavioral strategies (from micro-cellular to societal) that such systems constantly seek out in order to survive and reproduce in their environments (Thelen & Smith 1994; Thelen et al. 2001).

So, from this we believe it may be a bit of a straw-man to compare a culturalist position to simple nativist and empiricist positions as separate from one another. A more modern viewpoint (Lewontin et al. 1984; Oyama 1985) would need to view nativism and empiricism in a synthesized manner and conclude that the complex mutual feedback between the evolutionary and ontogenetic processes is what coordinates the development of categories. Therefore, the culturalist position is mainly innovative in that it posits a new complex adaptive system, that of language and culture use, as a third factor that plays a role in the feedback among levels to coordinate the development of categories.

The authors describe the culturalist position as "viewing language . . . as a complex adaptive system that is constantly coordinated by its users" (sect. 1, para. 8). However, they go on to indicate that they believe that a consensus needs to be reached on which approach, nativist, empiricist, or culturalist, is most appropriate in explaining the grounded development of categories, and therefore most useful for an engineer in developing a mechanism to implement robust category development in an artificial system. Even if language use and social interaction are shown to be another type of system that plays an important role in the development of categories, does this really mean that genetic evolution and/or individualistic learning would be shown to play lesser roles? No, we believe, and possibly the authors would agree. All three are involved, and understanding the development of categories necessitates understanding all three systems, as well as how they interact with and feedback on one another. As the authors say, the question is really one of levels of freedom, and which levels of adaptive systems are most involved in constraining which levels of freedom.

The authors maintain neutrality on the question of nativism, empiricism, and culturalism as to which, if any, theory best explains observations and data on human performance. However, they do state a position, saying that multiple sources of constraints are present in the formation of shared categories. They list three constraints: those coming from embodiment, those from the environment, and those from culture; and they generally identify nativists as emphasizing the first category, empiricists the second, and culturalists as throwing cultural constraints into the mix. It would seem that the true position they favor, and one we would very much agree with, is that all of these sources of constraints play important roles. Emphasizing one over the others always misses an important point, that it is the interaction between these constraints at different levels that is the key component of development. This article provides important results that will help us to tease apart the contributions of these various influences, using simplified models of developmental processes. However, we feel that the authors don't go far enough in pushing a synthesized view. Some people may still be stuck in a viewpoint maintaining the primacy of one type of constraint in the developmental process, but at least in terms of genetic and environmental constraints, it is clear to many that both play important roles, interact with each other, and the interactions between the mechanisms must be studied, as well as the mechanisms themselves.

Language and social interaction, as complex adaptive systems, would seem to occupy an intermediate level, in terms of time scale, between the relatively slow processes of evolutionary development, and the quick processes of ontogenetic learning. Therefore might they represent a kind of bridging level between the long-range and short-range processes? What level of social interaction is necessary so that a population will develop a shared set of grounded categories? The experiments in this article are a type of communication, but a very simple one at that. Is it really necessary to have a human-like language, or is some much more simple type of social interaction capable of developing shared categories? For example, is simply the fact of animals being social, where they have to act together and coordinate behavior, enough to provide some type of simple semiotic symbols that would allow for the development of coordinated categories? If any type of social interaction is capable of producing shared categories, does a more full-blown human language accomplish something even more in constraining levels of freedom? What extra mileage might a human-like natural language add to the development of shared categories?

It is not evolution, but a better game would need a better agent

Christian Huyck and Ian Mitchell

Middlesex University, Hendon, London NW4 4BT, United Kingdom.
c.huyck@mdx.ac.uk http://www.cwa.mdx.ac.uk/chris
i.Mitchell@mdx.ac.uk http://www.cs.mdx.ac.uk/staffpages/ianm

Abstract: Steels & Belpaeme (S&B) refer to the neural plausibility and evolutionary plausibility of their algorithms. Although this is not central to their goal of effective artificial agents, their algorithms are not neurally or evolutionarily plausible. Their communication games are interesting, and more complex games would lead to more effective agents. However, the algorithms could be improved either by using standard subsymbolic algorithms or by algorithms that are really neurally or evolutionarily plausible.

We accept Steels & Belpaeme's (S&B) main point that communication can increase the overlap between conceptual representations of both human and artificial agents. This said, we find several faults with it, including their inconsistent use of arguments, and their poor usage of evolutionary algorithms. We also find two related areas that should be addressed: hierarchical categories and more complex games.

S&B play fast and loose with their overarching methodology. They are inspired by the main approaches to human categorisation, but are not constrained by these approaches. This is fine when they are making points about artificial agents, but they frequently make references to evolutionary, environmental, and neural arguments.

For example, they state that their evolution simulations are too slow: "the agents need . . . at least 400 years" (sect. 3.3). This is a legitimate argument against people learning categories by evolution, but it is probably not appropriate for other types of biological systems. For example, fruit flies could learn the category in 20 days. A better argument against humans learning categories by evolution is that a person born to one language group, but raised in a second, learns the second. However, this argument is entirely irrelevant to their main point about artificial agents.

They also describe the use of environmental stimuli (sect. 2.4.1). S&B go into depth about the environment and mushrooms, but, in fact, their stimuli are just a set of 3-tuples. Their simulations have very little to say about the environment.

S&B make frequent use of the word *language* (e.g., sect. 4). However, their simulations only use labels. This is clearly a different thing from what is typically referred to as language, which includes syntax, grammar, semantics, and pragmatics. At best, their simulations are dealing with a symbol-grounding problem.

We are sceptical of the biological plausibility of radial basis function networks, though individual neurons do seem to map reasonably well to neurons. Moreover, the system they model has a network topology and learning algorithm that does not seem biologically plausible.

We find real problems with the genetic evolution simulations. S&B say that a generation is formed by retaining the best half of the previous generation, and a single mutated copy of each. This