Chapter 49

CRISIS! WHAT CRISIS?

PIERRE POIRIER
Université du Québec à Montréal

A few years ago, a well-known philosopher of cognitive science, Tim van Gelder, had an existential crisis of sorts. The crisis was brought about by a cognitive scientist who told van Gelder that philosophers had been left off a conference program because the organizing committee felt they did not have anything to contribute to the discipline. What was it, in fact, van Gelder painfully asked himself, that philosophers, given their training and the methods they master, can contribute to cognitive science?

To answer his own question, and resolve his existential crisis, he came up with a number of roles that philosophers were well positioned to play within cognitive science. The papers in this section were all written by philosophers and, taken together, they illustrate the depth and extent of the possible interactions between philosophy and the cognitive sciences.

One of the roles identified by van Gelder is that of the Building Inspector. As the name suggests, the philosopher-inspector works to ensure that the theories and explanations constructed by cognitive scientists respect the basic building codes of proper theory construction. Claude Panaccio’s chapter focuses on one such norm derived from a traditional philosophical doctrine, indeed a doctrine that rose to prominence in the Middle Ages: nominalism. His point is not that cognitive scientists must submit to nominalist principles but that a minimal sensitivity to such principles, in particular the principle called Ockham’s Cleaver, will help prevent needless confusion.

In philosophy, ontology is the field concerned with “what there is” – with the basic furniture of the world, as it is sometimes put – and nominalism is the claim that only singular objects, or individuals, exist: the cup in front of me, those in my cupboard, those in yours, and so on for all the cups in the world, individually. For the same is true of all other singular objects: individual atoms, individual animals, individual cars, individual planets, stars, galaxies, and so on. When the thesis is stated in this way, cognitive scientists, at least those who have not been too infected by philosophical theorizing, will no doubt say: “Well, of course, there are only individual things in the world.” However, as soon as they start devising their theories, cognitive scientists get in trouble with nominalism and betray their basic endorsement of the doctrine. What is the concept [CUP]?

One cognitive scientist might say that [CUP] constitutes the necessary and sufficient conditions that pick up “cup-hood” or “cup-ness,” that property shared by all cups. But according to nominalism, “cup-hood” or “cup-ness” is not a thing that can be grasped by...
minds or shared by cups, for it does not exist in the world. Remember that ontology is
about *what there is* and nominalism claims that “cup-hood” and “cup-ness” are not there
to be grasped or shared. Mindful of the nominalist position on properties, another cog-
nitive scientist, this one inspired by the brain metaphor, might say that [CUP] is that
attractor (in the neural network activation space) that picks up the correlation of visual
features exemplified by cups. But again, there are no such things in the world as corre-
lations of visual features. There are individual objects with their visual features, but no
correlations of visual features to activate an attractor in a hidden-unit space.

Panaccio himself is a nominalist but his aim is not to sell nominalism (although his
chapter contains a short section defending the option on epistemological grounds).
Rather, he does two things in his chapter that concern us. First, he shows how a “nom-
inalist sensibility” can be a heuristic for cognitive scientists working on concepts. He
therefore introduces what, following Boler (1985), he calls *Ockham’s Cleaver*, a prin-
ciple that states that one should be careful not to confuse the properties of signifier – in
our case, mental representations of concepts – with the properties of what they signify
or are about – in our case, categories of individuals. He points to a number of confu-
sions in the literature created by a lack of sensitivity to the Cleaver principle.

Second, he states and defends two constraints that any nominalist theory of concepts
should strive to respect. The first constraint is that *the only possible referents of con-
cepts are individuals*. Some concepts refer to no individuals at all (e.g., [FLYING
HORSE]); some refer to one individual (e.g., [THE PRESENT QUEEN OF ENG-
LAND]); and some to several individuals (e.g., [CAR]). Mindful of this constraint, cog-
nitive scientists should refrain from taking the referents of the above-mentioned
concepts as, respectively, the empty set, the singleton containing the present queen of
England, and the set of all cars. Concepts should not be taken to refer to properties,
attributes, features, or sets. It might be argued that in the case of sets, non-nominalist
talk can easily be translated into a form that respects this constraint, but the nominalist
would reply that it is nevertheless better to stick to the principle and refrain from talk-
ing of sets as referents of concepts: such a regimen may well prevent confusion from
creeping in through the back door. Panaccio’s second constraint concerns the mental
representations of concepts. They, too, are individuals (e.g., individual brain states).
Hence, they are not the sort of things that can be public or shared, as some cognitive
scientists would have it.

Another role recognized by van Gelder for the philosopher of cognitive science is
that of *Pioneer*. Throughout history, philosopher-pioneers have been both admired and
mocked: admired when their work led to some new science, political system, or ethical
principle; mocked when, as may happen to all pioneers, they go astray, get lost, or van-
ish without a trace (think of all those pioneers who lost their lives in the search for the
Northwest Passage). Chris Eliasmith is one among a few philosopher-pioneers defend-
ing the view that the study of categorization, or semantics generally, must begin with
the neurocomputational study of the cognitive brain. Only time, and much work, will
tell if a *computationa cognition neuroscience of meaning* is a foolhardy enterprise or if
it turns out to be our first real scientific foothold in semantics. Whatever happens,
Eliasmith’s work (here and elsewhere) is surely one of the best attempts at finding the Northwest Passage to Naturalized Semantics.

It should be noted that that the expression “computational neuroscience” has two meanings (even computational neuroscientists get them confused). These meanings are not incompatible, to be sure, but one is *stronger* and *includes* the other. The weaker notion is common in science, generally, when one names a new discipline or method *computational so-and-so* (where “so-and-so” nowadays can almost be any scientific discipline imaginable – computational chemistry, computational physics, computational biology, computational meteorology, computational genetics, and so on). In this more common sense, the expression simply refers to the use of mathematical modeling and computer simulation to understand the set of phenomena relevant to that field. Unlike other scientific disciplines (except engineering), however, neuroscience also affords a stronger meaning, whereby mathematical modeling and computer simulation referred to in the weaker notion are used to understand the systems studied by neuroscience as computational devices. Unlike the atmosphere, as studied by computational meteorology, the brain as a whole – but also its components (various brains systems), their components (neural assemblies), and perhaps even their components (neurons) – is understood by proponents of this stronger view of computational neuroscience as a hierarchy of nested computational systems. The computational neuroscience of vision, for instance, will understand visual phenomena such as color or motion perception as computational phenomena, and understand the brain systems devoted to vision as computational systems devoted to generating these phenomena. To mark the difference between the two expressions, I shall henceforth capitalize the stronger notion. It is important to emphasize that these two senses of the expression have different epistemic statuses. The weaker notion simply reflects the plain methodological fact that, given the power of today’s computers, almost any physical phenomenon is amenable to computer simulation, even those that are very computationally complex (think of *computational fluid dynamics*). The epistemic status of the weaker notion is thus quite straightforward: it is an observed socio-historical fact regarding the growing use of a particular set of tools in today’s science. By contrast, the stronger notion is a bold, and some believe quite false, hypothesis about the nature of cognition according to which the brain is a hierarchy of computer systems. For that hypothesis to be true, some structures in the brain must be representations, that is, they must possess semantic properties. Computing is the controlled transformation of structures that *stand for* the arguments of a function into structures that *stand for* its value. If a semantics mapping brain structures onto the arguments and values of a function can be described, then these structures become representations, the neural stand-ins of these arguments and values. If it turns out, however, that no semantics can be given for brain structures, then Computational Neuroscience is false.

In that context, Eliasmith’s chapter plays a double role. The first role is explicitly stated by the author. He argues that, whereas traditional accounts of categorization have all addressed the phenomenon from the point of view of psychology or linguistics, it should now be tackled from the point of view of computational neuroscience (not capitalized). In that view, categorization, like neuronal integration, motion perception, and
movement coordination, is a neural phenomenon that can and should be mathematically modeled and simulated on computers. With the demise of the strong reading of functionalism in the philosophy of cognitive science, which entailed the absolute autonomy of the psychological level of explanation with respect to the neurological level (the latter responsible for “mere implementation” issues), no one will object to a computational neuroscience of categorization. There are, however, many ways in which this could be achieved and the one Eliasmith chooses reflects the second role of his chapter. He attempts to provide a computational neuroscience of categorization by providing a semantics for a class of transient neural states (activation of neuronal populations). Such a semantics would show that the neural states are representations and, thus, that Computational Neuroscience is well-founded.

A further role identified by van Gelder for the philosopher of cognitive science was that of Cartographer. The philosopher–cartographer aims to show how the various theories, explanations and hypotheses that make up cognitive science, or indeed science in general, hang together (or not). This is what Edouard Machery and Luc Faucher do. They explore how two sets of theories of categorization, that is, social constructionist and cognitive evolutionary theories of categorization, can be integrated into a coherent framework.

To do this, they address a phenomenon that has largely been neglected by cognitive scientists working on categorization: the fact that some typologies (ways of dividing a domain into categories) seem more natural than others. According to most accounts of categorization, or cognition generally, the phenomenon should derive from the culture or physical world one is brought up in (or some mix of the two). Importantly, the phenomenon should not derive from any important fact about cognition. Accordingly, its study should not reveal anything important about the mind and hence cognitive scientists should ignore the phenomenon. With the rise of the evolutionary attitude in cognitive science, however, the phenomenon takes on a new urgency since it may reveal important fractures in our cognitive system: some typologies seem more natural because they form the representational core on which a cognitive module, i.e., a domain-specific cognitive computational system, is built. By contrast, those typologies that do not seem natural are this way because they are learned by some slow domain-general system or, worse, they must co-opt some module’s own representational core.

Instead of tackling this important issue in the abstract, Machery and Faucher focus on one such typology (or set thereof): those that underlie racist thinking, that is, the concept [RACE], and various racial concepts such as [BLACK], [WHITE], [HUTU], [TUTSI], and so on. Most scientists today (and certainly all those reviewed by Machery and Faucher) agree that racial concepts have no biological reality, that the current human species (Homo sapiens sapiens) cannot be biologically divided into subcategories. The question, then, is where do such concepts come from? Until recently, racial categories had been ignored by cognitive scientists and were mainly studied by social psychologists, anthropologists, and philosophers, who viewed racial categories as social or cultural constructs that were self-fulfilling, transient, culture-specific, time-specific, and socially or politically motivated. The social constructivists have recently been opposed
by a group of cognitive scientists inspired by the theory of evolution who claim that racial concepts are evidence of a domain-specific cognitive module. Evolutionary psychologists agree that there is no race module and that racial concepts are rather a by-product of the functioning of some other module, the nature of which is still in dispute (Machery and Faucher survey the important contenders). The structure of the debate is clear: social constructivism on one side and a number of evolutionary psychological theories on the other. As cartographers, Machery and Faucher weigh the merits and defects of each, ending their paper with a number of requisites for future theories of racialism that integrate the best contributions from each of the traditional adversaries.

Although they open their chapter by stating that they will address the fact that some ways of dividing a domain into categories seem more natural than others, they do not return to this issue, but stay focused instead on the categories that serve for illustrative purposes. It would have been nice if they had taken the time to state that position properly: Why do some ways of dividing the world seem more natural than others? Do all categories behave the way racial concepts do? Of course, one can read between the lines and infer to a great extent what their probable position is. Social constructionists presuppose that racial concepts are learned by some domain-general cognitive system. Evolutionary psychologists believe that racial concepts are derived through the activity of an innate cognitive module. Neither is completely right or wrong. Racial concepts are created in the interplay between, on the one hand, domain-general mechanisms that incorporate culturally constructed (and even perhaps culturally evolved) factors into the mind and, on the other hand, domain-specific mechanisms that, although they were not selected to build racial concepts, favor and constrain the formation of such concepts. As promising as this story sounds, it still needs a mechanism to generate the proper interactions between the domain-general and domain-specific sources of concepts.

Van Gelder’s crisis made him seek out what philosophers could contribute to cognitive science. His crisis may have been worsened by the fact that, over the past 50 years, philosophers have borrowed heavily from cognitive science: in the philosophy of mind, obviously, but also in the philosophy of language, in epistemology, ethics, aesthetics, and so on. In recent years, some authors have also suggested that philosophers could draw on cognitive science to help revive a traditional method of theirs: conceptual analysis.

In the first half of the twentieth century, it was believed that philosophy was not in the business of making claims about the world, but that it should rather concern itself with the clarification of our claims about the world. Our conceptual system, it was believed, is such that, even with the best of intentions, we are sometimes led to make meaningless claims or ask meaningless questions about the world. Philosophy was thus given the job of seeking out those conceptual errors that allowed meaningfulness to infiltrate into thought. One way to clarify thought was to translate the statements that express it into first-order predicate logic. Another important way to achieve this goal was conceptual analysis: analyzing the concepts that were thought to be responsible for such errors. This school, quite influential in the English-speaking world, was called Analytic Philosophy.
Conceptual analysis works by analyzing a concept into its constituent concepts. The statements resulting from such work were said to be analytic, that is, true or false solely on the basis of the meaning of their constituent concepts. And analytic statements were taken to be a priori (true or false independently of experience) and necessary (true or false in all possible worlds that contained them if true or false in the actual world). Analytical Philosophy is no longer with us in the form it had in the first half of the twentieth century: it received its coup de grâce when philosophers came to think of conceptual analysis as ill-conceived, that is, they came to believe that its intended product, analytical sentences, simply did not exist (not, at least, as they were understood and not, more importantly, as they had to be for the philosophers to do the work Analytical Philosophy expected of them): no sentence, Willard Quine showed in a famous paper, can be true or false solely by virtue of the meanings of its constituent concepts.

They say that everything old is new again and, as Elisabetta Lalumera reminds us in her chapter, conceptual analysis is currently undergoing something of a revival. Notwithstanding Quine’s pronouncements about the existence of the kind of analytical sentences that were thought to be necessary if Analytical Philosophy was to be in the thought-clarification business, the fact remains that some statements do intuitively appear to all competent speakers of a language to be analytical and necessary (the classical example being Bachelors are unmarried). What is more, people do intuitively confer a certain privileged epistemic status upon such statements. Facts such as these surely deserve an explanation. As Lalumera notes, there are two attitudes that naturalistically minded philosophers, that is, philosophers who want their explanations to be consistent with the best science of the day, may take toward this fact. She calls these two attitudes the inward and the outward approach. Proponents of the inward approach believe that the intuitions to be explained reflect something about the structure of our cognitive system. If that is true, then an important part of the explanation will be provided by cognitive science. Proponents of the outward approach believe that these facts reflect something about the structure of the world. After raising some problems with the inward approach, Lalumera defends her own brand of the outward approach whereby “relations among concepts, whether innate or acquired, mirror the relations among the real-world properties they refer to” (p. 1). Note that, as stated, Lalumera’s position conflicts with Panaccio’s nominalism, according to which there are no properties in the world, let alone relations between properties, for concepts to mirror. This, perhaps, illustrates one last prominent role of philosophers: disagreeing with each other. Whether or not readers are convinced by the arguments contained in these four papers, I believe that they cannot fail to be impressed by the role philosophers can play in cognitive science.
**Cohen-Handbook of Categorization in Cognitive Science**

### Queries and / or remarks

<table>
<thead>
<tr>
<th>Query No</th>
<th>Details required</th>
<th>Author's response</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No queries.</td>
<td></td>
</tr>
</tbody>
</table>