Epistemological problems in Cognitive Archaeology: an anti-relativistic proposal towards methodological uniformity

Duilio Garofoli\textsuperscript{1} & Miriam Noël Haidle\textsuperscript{1,2}

\textsuperscript{1} Zentrum für Naturwissenschaftliche Archäologie, Eberhard Karls Universität Tübingen, Rümelinstr. 23, 72070 Tübingen, Germany  
e-mail: duilio.garofoli@uni-tuebingen.de

\textsuperscript{2} Heidelberg Academy of Sciences and Humanities, Research Group “The Role of Culture in Early Expansions of Humans,” Research Institute Senckenberg, Senckenberganlage 25, D-60325 Frankfurt/Main, Germany

Summary - Cognitive archaeology (CA) has an inherent and major problem. The coupling between extinct minds, brains and behaviors cannot be investigated in a laboratory. Without direct testability, there is a risk that theories in CA will remain merely subjective opinions in which “anything goes”. To counter this risk, opponents of relativism originally argued that CA should adopt a method of validation based on “indirectly” testing inferences from the archaeological record. In this paper, we will offer a two-part analysis. In the first part, we will discuss problems with the original anti-relativistic agenda. While we agree with the necessity of developing a rational methodology for this discipline, in our view previous analyses have significant weak points that need to be strengthened. In particular, we will propose that “indirect testability” should be superseded by a methodology based upon deductive mappings from networks of theories, followed by a plausibility-selection stage. This methodology will be implemented by adopting an extension of Barnard’s (2010b) proposals for mapping hierarchical systems. In the second part, we will compare our methods with those currently adopted in the CA debate. From this analysis, it will emerge that some proposals in CA are inconsistent with our methodology and are incommensurable with those that are consistent with it. Furthermore, we will show that theories in CA can advance contradictory conclusions precisely because they have been developed using different methods. We conclude that a universal methodology, like that proposed here, is needed for CA to become more objective. It is also crucial for creating conditions for coherent and productive debate among different schools of thought in the field of cognitive evolution.

Keywords - Cognitive archaeology, Epistemology, Incommensurability, Theoretical Mapping.

Introduction

Twenty years ago cognitive archaeology (CA) emerged out of the Processual school and sought to distance itself from the subjective/interpretative approach adopted by the post-processualists. Interpretations were considered to be explanations open to manipulation that would have served to align an analysis of the archaeological record to the authors’ views. The interpretations could not be evaluated for how well they account for the target phenomenon on any objective criteria. Indeed, analyses were limited to \textit{ad hominem} and “inside” methods. In order to be able to interpret the archaeological record, investigators typically relied on their own subjective experience to situate themselves within the envisioned physical or social contexts of distant pasts, (Hodder, 1984, 1986, 1987; Johnson, 1999, Chapter 7; Trigger, 2007, Chapter 9-10; Shanks
The pioneers of cognitive archaeology criticized this approach as a form of opinion, based only on personal likes – it was “as wished for” archaeology (Renfrew, 1989, 1994; Bell, 1987, 1991, 1994a). The shared aim of these pioneers was to produce a methodology, with a clearly defined set of rules, to enable assertions about prehistoric cognition to be systematically tested.

The original epistemological objectives gradually faded into the background as new theories, deeply different in form and content, were developed to provide explanations about the evolution of mind and the emergence of behavioral practices considered unique to modern humans. Several frameworks, ranging from evolutionary psychology (Mithen, 1996) through cognitive neuroscience (Coolidge & Wynn, 2005; Wynn et al., 2009; Wynn & Coolidge, 2011) to computational theory (Barnard et al., 2007; Barnard, 2010a), were proposed to account for the properties of human mind and behavior (see Davidson, 2010 for review). These approaches, while grounded in different specific conceptions of the mind, nonetheless shared a common concern with providing cognitive/biological mechanisms underlying behavioral enhancements. A quite distinct tradition, rooted in the archaeological domain, sought to explain the same enhancements purely on the basis of socio-cultural interactions between individuals, without reference to biological constructs (de’Errico & Stringer, 2011; d’Errico, 2003; Hovers & Belfer-Cohen, 2006; Zilhao et al., 2010). This archaeological tradition uses as evidence for its position the presence of behavioral practices commonly associated with Upper Paleolithic populations in the artifactual record of early modern humans in Africa, as well as non-modern populations in Europe.

It is clear that these different theoretical proposals are not simply variations within a single school of thought. They range across many specific strands of argument using distinct approaches and methods that emanate from different communities of practice. Within this broader intellectual landscape, new sources of relativism continue to flourish, perhaps implicitly and less evident than in the past. This threatens to impede progress towards the emergence of a rigorous discipline of CA with a unified and coherent community of practice.

In this paper we propose a revision to the original anti-relativistic agenda, updating it to address the new epistemological challenges that have emerged over the twenty-year lifespan of CA. Over this lifespan the intellectual landscape has benefited substantially from rich and varied contributions from many disciplines. As we show later, arguments often navigate a specific course through intricate networks of related but qualitatively different theories.

As a fundamental assumption, we shall take for granted familiarity with the original anti-relativistic agenda (Bell, 1987, 1991, 1994a,b; Binford, 1987; Renfrew, 1994). Our focus will be upon selecting points in the context of presenting a novel, comprehensive methodology and space precludes an extensive discussion of the full range of issues associated with subjective interpretations. Moreover, in the meta-epistemological debate, a radical argument against emotionally driven approaches has been championed by proponents of the objectivist movement (Rand, 1964, 1967; Peikoff, 1981). In particular, these authors claimed that emotions arise as by-products of the process of value-formation. In consequence, they cannot be used to assess the validity of those same values. Objectivists reached the conclusion that emotions are irrelevant for establishing whether judgments are true or false. Clearly, if we accept that the only means of validation are individual likes and tastes, then the whole idea of “convincing” people of the validity of one theory over another loses any meaningful foundation. Acceptance of theories would be reduced to, for example, just counting how many people supported a particular theory’s contents. In contrast, we agree with objectivists’ view that there is simply no need to surrender to this pessimistic position: reason, indeed, can deliver a reliable method for validating theories and this represents the general goal of epistemological research. Our revision of the anti-relativistic agenda will be grounded, therefore, on the assumption that, contra the post-processualist
school, rational criteria can be provided to validate theories about extinct minds, brains and behaviors. Against this background, we will start by exploring the logical foundations required to establish a methodologically coherent CA. In particular, we will demonstrate the limitations of approaches that aim to provide explanations for the properties of the mind by drawing direct and unidirectional inferences from the archaeological record. Instead, we will discuss the need for CA to embrace a deductive perspective, which can allow networks of theories, constructed and mapped between multidisciplinary domains, to provide explanations of extinct minds’ properties. Since we cannot bring extinct minds into a laboratory, we will suggest that the concept of empirical testability introduced by Bell (1994a) is insufficient. For example, we will later show that symbolic thought cannot be inferred from the simple presence of beads or pigments in the archaeological record. Such evidence needs to be augmented with criteria of plausibility and logical validation to create an efficient strategy for selecting more viable theories from among less viable ones.

With foundational rules for a CA methodology in place, we will then explore the epistemological problems that can potentially confound meaningful comparisons of theories. Case studies will be used to demonstrate the actual existence of these problems in current key debates within CA. Specifically, we will elaborate how our proposed methodological framework provides the conditions and deep structure for what should hold for the proper and meaningful comparison of alternative theories. This framework has to use a theoretical vocabulary of considerable range and precision and this terminology is comprehensively specified in the Glossary at the end of the paper.

**How to get from artifacts to a theory of mind?**

The formulation of a universal method for CA is an ambitious and intricate problem to address. It implies a fundamental premise about the logical operations that need to be adopted to establish coherent connections between extinct brains, minds and artifacts in the record. An approach that aims to account for this problem must necessarily deal with multiple sources of data and types of theory. These need to be mapped one to another in a manner that supports justified inferences. Neural, mental and behavioral systems are inter-related entities that exert reciprocal explanatory influences one upon another. Although other system levels, such as social-cultural or bodily systems, are clearly relevant for CA, we will focus on just three levels to illustrate key points in our argument. Figure 1 shows the network of inductive, deductive and explanatory inferences linking interdisciplinary data and theory within and among these three qualitative distinct systems.

So, for example, research in neuroimaging, lesion studies, psychophysiology and neurobiology using extant species enables the induction of theories of neural systems. Similarly, experimental
Cognitive Archaeology

research in the field of cognitive psychology and comparative cognition, for instance, can be used to build increasingly more abstract theories of how the mind works, while data from cultural anthropology, ethnography and ethology can do the same with a theory of behavior. At the same time, the suite of theories makes it possible to draw deductive inferences that allow new data to be categorized and explained.

While most researchers would acknowledge that a logical framework of this type is required within CA, the literature itself is replete with examples of inferences that short-circuit the requirements it implies. Some classes of problematic inferences are highlighted schematically in Figure 2. Dashed lines represent inductive inferences that should not be made from archaeological data to minds and brains, while the boldened arrow highlights inappropriate inductions from archaeological data to behavioral theory. The overall organization here is crucial. Archaeology is limited to the behavioral domain and this prohibits direct inductive inferences. Given that certain forms of inductive inference are problematic within a single behavioral system without making inappropriate assumptions, then the problems compound when seeking to make inductive inferences from one system level to another.

The properties of artifacts in the record can only fully be explained “by means of” a theory of behavior, which is built in the present and mapped deductively on the past and can be tested. Handaxes have been argued to fulfill many functions, for example, in butchery (Mitchell, 1995), sexual selection (Kohn & Mithen, 1999), as weapons (Samson, 2006) or merely as by-products of manufacturing flakes (Davidson & Noble, 1993). Handaxes could be described as butchery tools because they have some necessary and sufficient properties that allow us to include them in a theoretical category that defines how a butchery tool ought to be. Inferences about how extinct species behaved with them require theories about the value of particular properties such as sharp edges, the presence or absence of wear, symmetry, the practicability of handling them safely or the contexts in which they were uncovered. Theories provide the necessary scaffolding for inferences and for these examples the properties in focus pertain only to behavior, we have deliberately made no reference to properties of mental or neural systems. This point applies only to inductions within behavioral systems; the problems are more acute in terms of what is required to move among system levels and this problem will be addressed later.

It will be recalled from the previous section that Bell’s candidate solution to the problem posed by the post-processualists was to provide a universal method for CA that would indeed enable the validity of claims about prehistoric minds to be directly tested on the artifactual record. It is

![Fig. 2 - Problematic inductive inferences within and between systems.](image-url)
clear that Bell thought that he could empirically validate properties of minds from the record. In doing so he conflated theories of behavior with theories of the mind. In the next section we will illustrate this form of conflation in detail for the case of artifacts for measuring weight and inferences about mental constructs. It should be clear, from the connective links in Figure 2, that empirical testability within archaeology is limited to the confirmation of behavioral hypotheses that are deductively constructed from theory of behavior and can be tested in the archaeological record. No empirical proof can be provided to infer directly from artifacts to the properties of the mind. Even if, as Bell implicitly argued, the behavioral system is constructed by a proper deduction from theory of behavior and tested empirically on the artifactual record, properties of behavior are not sufficient alone to explain the properties of the mind connected with it. Even the more solid empirical support of statements in the artifactual record cannot be used to test the characteristics of the extinct minds connected to them.

It is well known that a given pattern of behavior can be open to explanation by alternative theories, and this is obviously true for extinct as well as extant minds. Theory necessarily represents the starting point for testing explanations about the properties of the extinct mind. Some scholars have routinely drawn inferences from isolated parts of the artifactual record to specific characteristics of the mind, adopting a notion of empirical testability, where, for example, the presence of flower pollen in a grave is taken to be indicative not only of ritual behavior but a mind able to process symbols. To overcome these two problems we will introduce a holistic perspective for inductive, deductive and explanatory inference. Networks of inter-related theories of cognition, behavior and neural systems are mapped onto the artifactual record in order to explain properties of extinct minds and brains that cannot be tested through explicit behavioral experimentation.

Figure 3 focuses on mental and behavioral systems to indicate the general structure and functional rules for the method that we are going to propose for CA. This summarizes the arguments we have already introduced while also anticipating the content of what follows. In contrast to Figures 1 and 2, this new diagram highlights three features. First, it makes prominent the role of deduction from theories to properties of minds and of artifacts. Second, within this
bigger picture it is important to note that a key role is preserved for induction and deduction within research on mental and behavioral systems. Rather, the connection is mediated indirectly from properties of artifacts via theories of behavior through theories of minds to specific properties of minds.

This schema alone does not cleanly resolve the problem of providing objective criteria for empirical validation of theories – there is no escape from the problem of not being able to directly test explanations of mental systems that no longer exist. However, a proposal for non-empirical validation of theories will be advanced that arguably provides a rigorous framework for selection that is sufficient to overcome relativism.

Problem of testability

In the second chapter of the seminal book “The Ancient Mind”, James A. Bell (1994a) argued that archaeologists should aim to construct testable theories of prehistoric cognition as opposed to interpretations. According to his view, theories must be constituted by statements emerging directly from observation of the artifactual record and that, in turn, can be empirically tested on it. At the same time, other statements might be derived, which are not directly connected to the artifactual record, but can be subject to empirical refutation by virtue of logical connections with assertions that can be conversely tested on the archaeological evidence.

To help understand this logic, Bell revised the famous Renfrew’s (1982) example of the stone-based system of weights from the four thousand year old site of Mohenjo-Daro, in the Indus valley. By finding evidence of the existence of cubic stones, whose weights are multiples of the same unit, Renfrew advanced a theory of the cognitive processing of weight systems in this civilization. A relevant part of his argument can be structured in the following way (where <-> = implies):

a) discovery of calibrated stones <-> concept of weight + existence of units
b) existence of units <-> existence of modular measure
c) hierarchy of units <-> existence of a system of numeration

In other words, the discovery of calibrated stones suggests both the existence of a concept of weight and of constituent units (a), which are necessary to support the idea of modular measure, (b). Moreover, the existence of a system of numeration, organized in numerical categories, follows from the assessment of the hierarchical relationship among units. Therefore, the assertion of calibration is directly testable on the artifactual record, by seeking for similar stones that hold no weight relations and can therefore invalidate the inference in (a). Conversely, the assertions (b) and (c), are not refutable from a direct observation of the archaeological evidence (to this goal there ought to be written material from this society that documents the use of system of numeration, for instance), but they can be in any case rejected by considering their logical connection to (a), which is both testable and necessary to support the validity of (b) and (c).

This indirect approach to testability is formalized by adopting the concept of entailment, which Bell (1994a, p. 19) refers to with the expression:

“if X entails Y, then a mistake in Y indicates that there is a mistake in X”.

Therefore, in order to produce an entirely testable theory, statements not directly testable themselves must entail statements which are directly testable, so that the absence of clear evidence of a system of units in Renfrew’s example necessarily would lead to the invalidation of any potential conclusion concerning the existence of numerical systems. Indeed, the logic of entailment implies the obvious rationale that there is no concept of weight without the identification of artifacts that can fulfill the role of units, as well as no system of numeration.

In the previous section, we discussed the difficulties associated with empirical testing and
inference (Figs. 2, 3). In the case of Renfrew’s example of the Mohenjo-Daro weight system, the argument is constructed through a proper deduction from theory of behavior, which enables us to understand how a system of weights should be universally constituted. However, the fundamental problem here lies in the deceptive conviction that statements about prehistoric cognition are being tested, whilst actually testing assumptions applying within the behavioral system. As seems to be the case with much of the first wave of CA (see Renfrew, 1994; Preucel, 2007), both Renfrew and Bell focused on relationships that connect several components of a behavioral pattern, namely the stones and their potentially calibrated weights, in order to categorize the use of units of measure in a system of numeration. Classifying the use of calibrated stones with the label “system of weight” tells us little about the nature of the mental processes required to use it and the concept of weight that stems from this analysis can be different from the one we currently hold. Indeed, when referring to the single behavioral system, testability is warranted by direct inference from the archaeological record or, as illustrated above, through the logic of entailment. It is not by chance that Bell more than once used the expression “indirect way for empirical refutation of statements” (our emphasis). On the other hand, when mental processes are considered, the shift from a behavioral system to a mental one cannot be informed by empirical analysis alone, either directly or by entailment. For example, it is possible to characterize the behavioral practices in different hominid species with complex logical maps (Haidle, 2009, 2010, 2012; Lombard & Haidle 2012). However, here the behaviorist fallacy holds. You cannot infer the identity of underlying cognitive processes from observation of the behavior alone, even with an abstract schema of the sort used in our illustration of a system of numeration. Therefore, specific behavioral practices, considered in isolation from the rest of the behavioral architecture that characterizes one species/population, might be sustained by a different pattern of cognitive operations and ultimately associated with different mental capabilities.

This general point was originally discussed in the context of Wynn’s seminal review of cognitive evolution (2002). In his comments on that review, Deregowski argued that rotation of tri-dimensional figures and estimation of symmetry were not necessary to produce Early Acheulean handaxes. An alternative, but easier strategy can give the same result. In this case, a simple mechanism of perceptual priming would automatically have led to choosing the correct shape when presented with two possible alternatives. The hard epistemological challenge for CA is to put the flesh on the mere bones provided in Figure 3 to answer the question “how is it possible to test assertions about the evolution of the human mind from the archaeological record?” In framing an answer to this question, it is necessary to deal with the post-processualist counter-arguments to Bell’s agenda.

Given what has been argued so far, CA must chart a new and clear epistemology to avoid the black and white choice of two wholly unproductive options. One option remains Lewis Binford’s (1987) materialism, according to which drawing psychological inferences from material facts, via “paleopsychology” (Binford, 1965), is of dubious value. The other option is relativism, where anything goes, according to the authors’ interpretations, likes and so forth. In the next section we will explore a new epistemology offering some precise conceptual scaffolding for the “logical”, rather than empirical, validation of explanations.

The deductive approach

Structure of the deductive method

Our proposal to develop a new epistemology focuses on use of a deductive framework to explain the archaeological record. This deductive method aims to map a suite of theories that account for systems of interest and how they behave. Behavioral, mental and neural systems are causally inter-related, each with their own qualitatively distinct architectural properties, and exert reciprocal constraints one upon
another. As we explain later, in order to better understand these reciprocal influences, we need a well defined set of macro-and micro-theories of how each type of system works (Barnard et al., 2000). The notions of macro and micro-theory are deeply rooted in a hierarchical conception of how assembled systems are constituted and serve to explain how entities that are part of each system behave (Newell, 2000). A macro-theory can be roughly defined, in fact, as a theory that explains how the subcomponents of an assembled system interact, while each of these subcomponents requires in turn a micro-theory to explain the properties of its constituent parts.

The challenge is further complicated by the fact that our focus in CA is on understanding the neural, mental, and behavioral systems of not just one species of hominoids, but many. Because multiple macro- and micro-theories of qualitatively different systems and many species are required, we clearly need a way of thinking about all these inter-relationships in a systematic rather than piecemeal way. To realize our method, we have to organize the deductive framework and specify the first major premise that can lead to general principles of interaction that govern the behavior of any system. Figure 4 illustrates this by referencing two new constructs: Meta-theories that relate system levels for a given species (e.g. Homo sapiens sapiens and Pan paniscus), and the concept of Core-Theory which aims to capture what governs interactions among system levels for all species across any and all system states. The later construct and its detailed properties will be incrementally built up as our argument develops. For the present, it is sufficient to note what Figure 4 highlights is that Core-Theory must be induced from Meta-Theories. In this paper, Figure 1 diagrammatically introduces the function of an explanatory inference and its more precise definition can now be stated. “Explanatory inference” is the logical operation adopted to construct a particular Meta-Theory and it refers to the bidirectional explanatory power these
systems reciprocally hold. This procedure of horizontal mapping is repeated for different species/populations and many meta-theories are produced until all of these are aggregated in a synthetic “core-theory”.

The value of core-theory is that it should enable us to predict through explanatory inference, for instance, how the variation in the architecture of a mental system (M) affects a behavioral system (B), by virtue of the principles that universally bridge and constrain the various systems. However, a core-theory appears as a pure abstraction and should be considered as a way to account for the comparisons of meta-theories and the extraction of invariant rules of system behavior.

The next step for CA is to recruit the power of deduction to perform a vertical mapping from the level of a core-theory to a target level of interest (Fig. 5). In CA, only one system is available within the target level - the behavioral architecture found in the artifactual record, albeit where the evidence of the complete system is necessarily partial. The obvious point is that deduction is required to infer properties of the missing target mental and neural systems. This method, as depicted in the figure, can be represented with the following expression in formal logic:

\[
\text{If } X \leftrightarrow Y \leftrightarrow Z \text{ and } X = X^* \text{ then } X^* \leftrightarrow Y^* \leftrightarrow Z^*
\]

In line with the structure of deductive arguments in logic, the validity of the conclusion follows directly from true premises and for every given X it would be possible to find a specific Y and Z, by virtue of the universal rules that are implicitly stated in the major premise.

Figure 5 is a refinement of our earlier Figure 3 that now makes clear the role of explanatory inference. If a direct horizontal mapping is impossible from the target behavioral architecture to the other systems in the same level, the only way to fill the gaps of knowledge results in vertically inferring them by virtue of the universal principles of connection that bind elements in the accessible

---

**Fig. 5 - Vertical mapping from the core-level allows to complete the missing fields (interrogative marks) in the target level, in relation to the structure of the behavioral architecture in the target level.**
meta-theories and are synthesized by the core-theory. In other words, the diagram, as formulated, tells us that having different perspectives on how neurons, cognitive systems and hominoids behave, warrants sufficient generality to infer from them how a given behavioral architecture, detected in the artifactual record and linked to a target species/population, is associated with a mental and a neural architecture. “Implemented theory” is the way we define the result of the vertical mapping between the core-theory and the behavioral architecture in the target level. In other words, an implemented theory is the final body of theories that is provided in CA to explain properties of extinct minds and brains. A schematic synthesis of the whole logical process that defines our proposed methodology is shown in Figure 6.

This deductive method obviously relies on two assumptions. First, relationships between system levels are subject to deep abstract principles (a), which exist independently from the specific theories adopted to explain them. This means that variations in the properties of a system necessarily produce changes in related properties of connected systems. For example, capacity is a property of both the neural and the mental system that accounts for the amount of information a neural network, as well as a cognitive architecture, can potentially handle (Halford et al., 2007). According to (a), increasing the capacity of the neural system, for instance by altering the architecture of constituent networks, increases also the capacity of the cognitive one. Therefore, the behavioral repertoire that can be potentially handled by these enhanced systems is also “increased”. At the same time, our second assumption (b) implies that those principles applied in the same way in the ancient past as in the present. It is well known, as Hume (1739) argued, regularities in the present will not necessarily apply in the future, but we believe we can trust that biological principles past and present are congruent simply by virtue of the evolutionary relations that link organisms. Our proposed methodology shares with Bell’s original conception the property of assuming indirect strategies as a tool to validate theories. The main difference between our approaches, however, lies in the fact that we do not recognize the logic of entailment as sufficient to check the validity of statements between two different aggregated systems.

As an illustration of the difference between using our holistic and deductive approach as opposed to that relying on entailment, we can consider Acheulean tools. These are created by the staged removal of flakes from a core to create a bifacial and symmetrical entity. Wynn (2002) argued that mental rotation of three-dimensional figures played a key role in the process of manufacturing these tools. However, nothing in the record alone can directly be used to argue against the involvement of mental rotation without a theoretical context to support an appropriate deduction. In contrast, implications can be invalidated in cases such as Renfrew’s example, when stones with non-standardized weights are found. The point is that inferences from evidence across different domains, when isolated from a theoretical context, can lead to the formulation of distorted conclusions. In fact, statement X “there are symmetric relationships between parts of this artifact” entails statement Y “symmetry is produced by tridimensional rotation of figures in working memory”, but there is no empirical evidence that can disprove this entailment in the artifactual record, because the mind does not fossilize.
Problems like these need to be solved only by checking if the whole behavioral architecture of the population that produced those artifacts is compatible with a hypothesized mental architecture. This needs to be derived via a deduction from the core-theory that is built upon the firmer foundation of empirically warranted meta-level theories. We need to go beyond the idea of testing single statements and the adoption of linear chains of inferences characteristic of Bell’s falsificationist agenda. Instead, we need to shift to an alternative conception of “testability”, more similar to a form of holistic evaluation of theories (Godfrey-Smith, 2003, p. 31). To reconstruct missing mental and neural systems we need to examine not single statements that are entailed, but entire meta-theories. For a given target behavioral architecture, a core-theory of system behavior would predict properties of mental and neural architectures (i.e.: the core-theory that explains the connection among systems in modern humans and apes entails the conclusions about the mental architectures of other extinct hominoids).

In framing our approach so far we have concentrated on only three systems. This should not be taken to exclude contributions from macro-theories of other relevant but qualitatively distinct systems. Clearly, properties of the biomechanical and sensory systems of a particular body morphology also constrain relationships to cognitive and neural systems. Such constraints may be informed by models of embodied cognition (Lakoff & Johnson, 1999; Clark, 2008) and their application to the study of material culture (Malafouris, 2008a,b,c, 2010a,b; Mithen & Parsons, 2008). The same applies for constraints that propagate from higher order social systems. An account of demographic changes would give more power to deductive mapping, both in the horizontal and vertical component. One example of this could be the hypothesized link between group size and rates of cultural innovation, which comes with at least some authority from mathematical modeling (Shennan, 2001; Powell et al., 2009). Furthermore, recent empirical analyses have proved the existence of a correlation between technological complexity and population size. In particular, Kline & Boyd (2010) meta-analyzed a set of ethnographic data on artifacts and behavioral practices of populations living in different Oceania islands and concluded that islands with small populations had a simpler marine foraging technology (but see Read, 2012 for a counterargument). More widely, an extension of the three-systems logic would offer additional constraints and allow a more informative mapping to the archaeological evidence (i.e.: both theories of the body and demography are informed by paleoanthropology).

In a similar vein, constraints on neural architecture can come from paleoneurological studies on endocasts and neontological enquiries characterizing relationships between variation of brain size, shape and functions in human and non-human primates. The paleoneurologist Emiliano Bruner, for instance, has demonstrated that modern humans, but not highly encephalized Neanderthals, are characterized by a non-allometric expansion of the upper parietal regions of the endocast (Bruner, 2003, 2004, 2008, 2010, 2011), an autapomorphic trait that is supported by a specific ontogenetic phase in modern humans that has no counterpart in species that predate anatomically modern humans (Gunz et al., 2010; Neubauer et al., 2010). This morphological variation, which determines the characteristic globular form of the modern human brain (Lieberman et al., 2002), contrasts with the elongated structures of more archaic populations. This difference has been associated with potential disproportional white matter expansion and enhancement of parieto-frontal connectivity (Bruner, 2003, 2004; Coolidge & Wynn, 2008). This enhanced connectivity, in turn, could be a candidate biological substrate for selective cognitive advantages (Jung & Haier, 2007) in modern humans. This hypothesis has been recently tested in neontological studies on modern humans. A slight correlation between brain globularity and information processing speed has been demonstrated (Bruner et al., 2011). Additionally, these researchers claimed that this effect might have been more pronounced on an evolutionary time-scale. This
form of analysis well illustrates how empirically supported constraints relating system levels can contribute to the substance of core-theory.

One key question remains to be addressed. Under circumstances where there is no direct empirical evidence to refute theory, and a number of implemented theories are available, how do we select the candidate that best accounts for the context of the record?

Problem of validation: criteria for selection

Any deductive argument can lead to false conclusions if the premises are false (Godfrey-Smith, 2003, p. 41). Use of our methodology involves a series of stages (Figs. 5, 6) and false conclusions could arise as a result of the adoption of flawed macro-theories, by performing erroneous horizontal mapping between systems, or by creating a corrupted core-theory. Even an adequate core-theory would allow errors to propagate into the vertical mapping, if its minor premise is wrong, as for instance in case where there is some problematic interpretation of the artifactual record, and hence flawed behavioral architectures. In this respect, our proposed methodology allows us to pinpoint with some clarity where intricate arguments can be flawed. This same property also means that we are going to need two different stages to decide among candidate implemented theories. These two stages once again correspond to the two main phases of mapping.

The first stage basically involves the construction of a core-theory. Here, evaluation simply means that principles must be supported by empirical data. Experimental work would serve to define macro-theories of the assembled systems of interest, by virtue of the study of their micro-theoretical components, and to understand the way they constrain each other. This validation stage is the “easier” part of the epistemological problem of CA. This should not be taken to mean that the construction of a core-theory is an easy task, but that in this phase there is still a connection to hard data. Extracting principles of interaction between levels of architecture for extant species is far from an easy task. There is not much agreement about how minds should be represented in whole or in part. Even though there are candidate unified theories of cognition (e.g. Newell, 1990) and macro-theories of mind (e.g. see Barnard et al., 2000), there are vast numbers of micro-theories to account for specific phenomena and not much in the way of consensus about how best to decompose minds or about how individual micro-theories can be reconciled as parts of a bigger picture. However adequate the micro-theories might be within a discipline, a key challenge for cognitive neuroscience, cognitive anthropology and comparative psychology would be to develop convincing rules to connect their own system-level concerns with those at an adjoining level - a task that becomes even harder as the number of constraints to be accommodated increases.

A crucial but also hard aspect of the validation process lies with the vertical mapping from core-theory to the target level “missing” system, be it mental or neural. As argued earlier, evidence cannot play a part here. The “logical validation” that is performed within our methodological schema requires criteria. Some criteria for selection can be borrowed from David Chalmers’ (1995, 1996) analysis of the “hard problem of consciousness” (for a review see Searle, 1997, 2004). This problem shares with CA the fundamental epistemological issue of inaccessibility of the object of science to scientific methods. The subjectivity of first person experience can no more be investigated with scientific methods, than can the minds of extinct hominoids. Chalmers’ solution involves a series of non-empirical principles to evaluate candidate theories of non-observable phenomena. The six criteria proposed by Chalmers are adapted below for CA:

a) Internal Coherence: when mapped from a core-theory to the target level, the principles that bind architectures must be preserved in the target level. In other words, the rules adopted to infer properties of the missing systems should be coherent with those predicted by the core-theory. Not complying with this criterion would imply an incoherent vertical mapping and the whole deductive process would be corrupted.
b) **Simplicity**: Theoretical schema should aim for simplicity in any of their component parts. In particular, the result of vertical mapping should be a simple target level meta-theory. However, it is worth noting that simplicity is not always possible, due to the fact that there are no guarantees that the world is simple (Godfrey-Smith, 2003, p. 55). It should be possible to validate complex theories as well, if the payoff exceeds the increase of complexity.

c) **Homogeneity**: Theories must show no gaps in aspects of reality that they intend to cover.

d) **Inclusivity**: Theories should not deliberately exclude aspects that cannot be accommodated within their logical structure.

e) **Consistency**: There should be no contradictory parts in a single implemented theory. A statement and its negation cannot be demonstrated within the same theoretical schema.

David Chalmers, in his book “The Conscious Mind” (1996 - Chapter 6), supports the idea that the six criteria for selection exert constraints on the plausibility of theories and that this can act as the gold standard for assay in non-empirical situations. However, Chalmers does not provide any precise definition for this concept. He prefers to use an example to discuss the idea that two alternative theories might be both perfectly rational in terms of the logical connections between elements, but they can hold a different level of plausibility. In this way, he imagines the situation of two competing hypotheses. According to the first one, the world has been created fifty years ago, together with all fossils and memories, while in the second hypothesis evolution really happened as we know it. Comparing these hypotheses, one must conclude that both are rationally conceivable, but the first one is implausible, because it is too complicated and based on a series of unnecessary assumptions.

To solve this series of issues, we propose a definition of plausibility that takes into account the number of *ad hoc* hypotheses that are produced to connect the core and the target level during the construction of an implemented theory. *Ad hoc* hypotheses, in fact, could be used to resolve issues with each of the principles derived from Chalmers (1996) and listed above. They can be employed to correct problems of coherence in the rules adopted to map the various elements within the whole theoretical schema (crit. a), to cover gaps of knowledge left behind or to deliberately rule out uncomfortable parts (crit. c-d). They may also be used to justify contradictory aspects of the theory by creating “exceptions” (e), while simplicity (b) results as deeply influenced by the number of changes that are produced to align with the other criteria. It follows that implemented theories that need to be adjusted in any of these ways are less plausible than those able to perform a coherent mapping without recourse to logical alterations. Therefore, this variant of plausibility works by embracing Occam’s razor. Selection among alternatives is now dictated within an eliminative perspective (Platt, 1964), that rules out all the more implausible alternatives *en route* to adopt a preferred theoretical schema. To help understand this generic logic, we can draw on Walton’s (2001) discussion of the case of a juridical diatribe in Ancient Greece. Two contenders, a big man and a little one, were involved in a fight without witnesses. Each of the contenders had to convince the jury that the other had started the fight. It would be implausible for the jury to think that the small man assaulted the bigger man first, without posing some *ad hoc* hypothesis to justify this inference. For instance, claiming that the little man is an expert martial artist would raise the level of plausibility of the hypothesis that he actually started the fight. Martial experience would compensate for lack of size. However, if it were not possible to empirically check the fact the little man was an expert fighter, then there would be no grounds to suppose that this hypothesis is actually more plausible than the first one and there would be therefore no reason to prefer it.

In CA, reference to evidence is typically more indirect and elusive, so that even the most plausible implemented theory would quite likely retain a speculative component in the vertical mapping phase of the framework. But as we have argued,
this does not result in epistemic relativism or equivalence in the value of the conclusions: selection for plausibility is grounded in explicit criteria for theory validation, where Bell’s empirical testability perspective is not applicable.

Comparing alternative theories

Problem of Incommensurability

To evaluate the content of alternative implemented theories, it is necessary that the same basic principles are applied while developing those theories, from micro-theories to the full implementation. Theories of complex systems are seldom simple and cannot be easily reduced to unitary elements that can be easily manipulated. As we shall see, a network of theories at different levels may be organized into hierarchical layers and, when one is mapped to the others, relationships within and among their constituent layers must be treated in a disciplined way (Barnard et al., 2000).

If everyone were to use a different set of rules for mapping among system levels, then the final cognitive archaeological theories that are implemented and compared would be incommensurable. The word “incommensurable” implies that the logical structures involved in the construction of their premises and in making deductions differ. Consequently, it would be impossible to properly compare those implemented theories. Incommensurability can be responsible for confusion and relativism, because some theories that are presented as real alternatives to other ones are alternatives not because they are theoretically advantageous, but because they have been assembled with an improper methodology. At the end of the paper we shall illustrate these problems of comparison with two case studies of inferences from the presence of beads found in Neanderthal and Early Modern settlements.

Problems of incommensurability can be further refined into three categories:

1) General Incommensurability: Meta-theories (see Fig. 4) are assembled with rules that are not universal, but specific to the theory being implemented. (i.e.: some parts of one system architecture are mapped onto different parts of other architectures, or parts are omitted, etc...). Here the mappings are incoherent because their premises are problematic.
2) Incommensurability within architectures: assembled systems are mapped correctly, following a universal logic, but the rules of construction within each mapped system are different, so that the system architectures themselves are incongruent. Incongruence, in this case, is the cause of incoherent mappings.
3) Incommensurability between architectures: assembled systems are constructed according to the same rules, but the different layers that constitute each architecture (Fig. 7) are again incoherently mapped to the layers in another architecture using a set of different rules. (Same system structures, different logic to connect them).

In order to attempt to find a solution to these three problems of incommensurability, we will refer to the model proposed by Philip Barnard and colleagues (2000) for mapping theories, revising it on the light of our methodological framework and suggesting its use as a potential uniform method in CA.

Barnard’s model: rules and problems

Barnard et al.’s framework was originally derived from research on human use of modern rather than ancient technological systems and therefore the problems it addressed parallel those of CA (see Barnard et al., 2000 and reference therein). A key objective of their work was to create a method to bridge systemic architectures of a qualitatively different nature (e.g. mental & technological systems) in a coherent way that would support an understanding of the conjoint behaviors of human and technological interactors.

The basic logic of their approach is based on hierarchical decomposition (Fig. 7). Every architecture is an assembled system [A], composed of a series of basic interactors [B]’s. At the same time,
each of these basic interactors is itself composed of constituent interactors [C]s. A macro-theory of an assembled system is required to explain how the basic units interact as parts of the overall system, while micro-theories are required to explain how the constituents of each basic interactor govern its behavior. So, each interactor is a behavioral entity that occupies a specific position within the hierarchy of an assembled system and is so named because it interacts with other elements within the same layer.

The behavior of any interactor is determined by two main considerations that act concurrently on it: (1) the nature of that interactor’s constituents, and (2) the constraints that are exerted on it by the other interactors present at the same level within the hierarchy. So, for example, according to Barnard (2010b), a behavioral system assembled to make stone tools might be composed of a human agent [B1], a hammer [B2] and a core [B3]. A tool using system might have the same human agent [B1] but now replace the core and the hammer with the tool [B4] and an animal carcass [B5]. Likewise, an hypothesized mental architecture of a Neanderthal might be decomposed into a particular set of mental subsystems (Bs) and that of a modern human into a different assembly of Bs, with some of these in common between the two architectures and others distinct. Similarly, their brains would be composed of neural circuitry with shared cellular electrochemistry but different network architectures whose differences really can be mapped onto hypothesized differences in mental architecture. Figure 8 now illustrates a schema for mapping networks of neural, mental and behavioral theories.

Assembled systems can be mapped horizontally by adopting the following basic principles:

1) An assembled system [A] in one layer (e.g.: here neural or mental) enables its collective capabilities to behave as a basic unit within a system assembled at some superordinate level, here mental or behavioral (A→B), where its properties now constrain the behavior of the new superordinate assembly.

2) The behavior of a basic unit [B], when incorporated into a superordinate system, must also carry with implications from a relevant body of micro-theory from the lower layer (B→C), since these also constrain how the relevant B may behave within the superordinate system. For example, for system assembled with an expert stone knapper [B] or an inexpert one [B’], the two may share a common macro-theory of how their minds are composed overall, but require different micro-theories of their perceptual, manual and planning skills.

3) The relationships are bidirectional, in the sense that the [B]s in a superordinate system can constrain an [A] in the subordinate one (B→A) too (but see below for an alternative conception of the principle of directionality). To continue with the same example, the particular micro-theory needed for an expert stone knapper would have arisen only if that individual had a history [383x39]www.isita-org.com [0x0]www.isita-org.com [428x642]21 [237x643]J. Hawks [195x643]D. Garofoli & M. N. Haidle [428x642]J. HawksD. Garofoli & M. N. Haidle [65x431]each of these basic interactors is itself composed of constituent interactors [C]s. A macro-theory of an assembled system is required to explain how the basic units interact as parts of the overall system, while micro-theories are required to explain how the constituents of each basic interactor govern its behavior. So, each interactor is a behavioral entity that occupies a specific position within the hierarchy of an assembled system and is so named because it interacts with other elements within the same layer.

The behavior of any interactor is determined by two main considerations that act concurrently on it: (1) the nature of that interactor’s constituents, and (2) the constraints that are exerted on it by the other interactors present at the same level within the hierarchy. So, for example, according to Barnard (2010b), a behavioral system assembled to make stone tools might be composed of a human agent [B1], a hammer [B2] and a core [B3]. A tool using system might have the same human agent [B1] but now replace the core and the hammer with the tool [B4] and an animal carcass [B5]. Likewise, an hypothesized mental architecture of a Neanderthal might be decomposed into a particular set of mental subsystems (Bs) and that of a modern human into a different assembly of Bs, with some of these in common between the two architectures and others distinct. Similarly, their brains would be composed of neural circuitry with shared cellular electrochemistry but different network architectures whose differences really can be mapped onto hypothesized differences in mental architecture. Figure 8 now illustrates a schema for mapping networks of neural, mental and behavioral theories.

Assembled systems can be mapped horizontally by adopting the following basic principles:

1) An assembled system [A] in one layer (e.g.: here neural or mental) enables its collective capabilities to behave as a basic unit within a system assembled at some superordinate level, here mental or behavioral (A→B), where its properties now constrain the behavior of the new superordinate assembly.

2) The behavior of a basic unit [B], when incorporated into a superordinate system, must also carry with implications from a relevant body of micro-theory from the lower layer (B→C), since these also constrain how the relevant B may behave within the superordinate system. For example, for system assembled with an expert stone knapper [B] or an inexpert one [B’], the two may share a common macro-theory of how their minds are composed overall, but require different micro-theories of their perceptual, manual and planning skills.

3) The relationships are bidirectional, in the sense that the [B]s in a superordinate system can constrain an [A] in the subordinate one (B→A) too (but see below for an alternative conception of the principle of directionality). To continue with the same example, the particular micro-theory needed for an expert stone knapper would have arisen only if that individual had a history
of training and practice in similar behavio-
oral architectures in the past.

Taken together, it should be clear that prop-
ties and behavior of interactors present in mental
and behavioral systems are a product of families of
constraints and that these tightly connected sys-
tems reciprocally influence each other. This idea
of bidirectional influence can be contrasted with a
more neurocentric perspective, where causal rela-
tionships between systems are frequently cast as
unidirectional and commonly oriented from the
neural to the mental right through to evidence
in behavioral systems. For example, Klein (1995,
to be the abrupt emergence of Upper Paleolithic
innovations by reference to a punctiform muta-
tion in modern human brain architecture, incurring
at ca 60 kya. In his view, this neural enhance-
ment promoted modern human cognitive capa-
bilities and in particular the extraordinary ability
of our species to innovate.

While Barnard’s framework has much to
commend it, there are only a few examples of
its application to practical cases in modern and
ancient technological systems. It is clear in broad
terms from Barnard’s description that the things
that interact in neural architecture relate to cir-
cuits, cells and their electrochemistry, those that
interact in mental architecture concern states of,
and operations on, “information”, while those in
behavioral architecture relate to changes of state
in animate and inanimate entities in our physi-
cal and social worlds. However, for each of these
systems there is a great deal of ambiguity about
what really constitutes a specific interactor and
which level it should be assigned to in a hier-
archy. This problem is particularly acute when
it comes to mental architecture. In the field of
cognitive psychology as a whole, there is lit-
tle agreement on how we might best define the
components (i.e.: [B]s) of the mind. Even worse,
there is a vast numbers of candidate micro-the-
tories applied to specific domains of mental life
among which it is hard to choose on the basis of
evidence currently available and little in the way
of a body of macro-theory to organize them. We
have a problem in determining what the [B]s and
[C]s are and what layers to assign them to when
we apply this approach to CA.

If the aim of the diagram is to provide a
tool that is universally valid, independently of
the macro-theories we choose to adopt, and the implemented theories produced with this model should be commensurable and accessible to the criteria for selection we proposed earlier, then it is necessary to generate a list of rules of functioning. These rules should be followed universally by all implemented theoretical proposals. If this fundamental condition is not met, the model once again risks allowing relativism. In this case allowing anything to be an A, a B or a C enables a theorist to propose connections that reinforce their specific theory and make it difficult to compare that theory with others.

By way of clarification, though necessary as a starting point, Barnard’s descriptions of the interactors in neural, mental and behavioral systems leave us with a problem of ambiguity to resolve. The main problem with the absence of precise rules for construction of architectures is that they might be built by adopting different strategies. As a consequence, arbitrary rules of construction can lead to architectures that are incongruent when mapped. It appears from Figure 8 that while mental architecture is defined as the entire structure of one species’ mind, constituted by a set of units that interact, which in turn can be decomposed in constituents, the neural architecture does not seem to account for the entire brain, but only for one extensive part of the whole circuitry. The behavioral architecture, as well, is focused on just one behavioral practice and the entities it involves. It does not seem to address the entire spectrum of practices accounted for by the mind of a particular hominine and the wider culture within it is embedded.

This problem with the original formulation has been recognized by Barnard (personal communication) and relates to the “Problem of optional incompleteness”. Complex biological systems, which involve huge networks of interactions, can hardly be depicted by synthetic diagrams. This can be seen in neural architecture, for instance, for which a complete description of the whole set of interactors that constitute it would require considering almost limitless amounts of entities and processes (e.g. \(10^{11}\) neurons plus many neurotransmitters, hormones and the endocrine system). The same applies to behavioral systems. Over the course of a human lifetime and over all the tasks accomplished by people in different human occupations and roles, there are equally vast numbers of possible behavioral systems. Following these assumptions, it appears that all layered architectures cannot be completely inclusive. However, Barnard notes (personal communication again) that his focal point for theory development is a macro-theory of mind that is fully specified, while his strategy is to leave adjacent layers only partially specified in a manner that most efficiently informs the development and testing of his theory.

In our view, for the purposes of developing this framework for CA, the problem of “optional incompleteness” must be taken into account by embracing a “realist” agenda that would allow us to recognize the limits of incompleteness. We should not confuse “optionality” with “anarchy”. This means that the optional choices in decomposing and configuring architectures must be made a priori. They need to be considered in the stage of epistemological discussion and not a posteriori, just before the application of a particular implemented theory. In other words, the fact that architectures cannot be completely inclusive by nature cannot be used as a justification to allow any potential manipulation of their structure and hierarchical organization.

Assembled systems need to be as congruent as possible, starting from the nature of the [A]s, which must be set at a comparable level of reference/complexity. Even though we might say that a chosen mental architecture is incomplete, this should not allow us to arbitrarily change the level of reference in the other architectures by replacing the brain with a part of it and then mapping the partial neural architecture on the whole mental one. So, if the [A] in the mental architecture is intended to represent the overall structure of the mind, then obviously we would perform an incoherent mapping if we chose some substructure such as the arcuate fasciculus’ circuitry to support this role in underlying neural architecture.

Now, examination of Barnard’s diagram reveals a conceptual difference in the
composition of mental and neural architectures on the one hand and behavioral ones on the other, that indicate we need to add to his schema to achieve full coherence of horizontal mapping. Also, this addition will help counteract inappropriate inferences from behavioral systems to mental ones.

With neural and mental architecture, the set of basic units that form the assembly are invariant – all “standard” humans have the same number of mental subsystems/processes and the same sets of basic neural circuits – although there will, of course, be naturally occurring exceptions. Variation in capability of the system, such as expertise at manual skill, vocal communication or problem solving, occurs in the level captured by micro-theory. The same clearly does not apply with behavioral architectures. Across, for example, Mousterian tool making, hunting, ornamental marking, food gathering and preparation, procreation, caring for the young and old, migrating, or conducting simple burial, there is significant variation in the Bs that enter into the systems and accordingly variation in both layers of the hierarchy for behavioral systems.

As expanded upon later, two interconnected points follow from this observation for how we should frame use of horizontal mappings. First, when making horizontal connections from mental to behavioral architecture, the selection of what is connected must be made in a manner that is not generic, but sensitive to the target context. Prior to the invention of writing, stories that could not be depicted had to be memorized and vocally transmitted and learned. Second, there is the inverse problem that when attempting to make horizontal connections the other way – from evidence in a behavioral system to mental architecture - a formulation such as Figure 8 may be used to justify erroneous generic inferences. For example, existence of a basic interactor in a single behavioral system involving a target species, such as bead ornamentations, has been used to argue for the generic presence of symbolic mental capabilities (Zilhao et al., 2010; d’Errico et al., 2005; d’Errico & Stirner, 2011). However, the wider context of other behavioral systems for that species may contain evidence that the presence of bead ornamentations is open to explanation by non-symbolic mental capabilities.

Barnard’s model extended

It is worth exploring in a little more detail the kinds of mapping from mental architecture to a behavioral architecture that can lead to distorted conclusions. Improperly constructing an architecture using a single example or class of behavior, such as ornamentation, rather than a larger “repertoire” of behaviors, represents a variant of the second type of incommensurability, that within architectures. In this case, the same macro-theory of mental architecture can be mapped onto different forms of behavioral architecture, resulting in markedly different implemented theories.

An example of this class of improper reasoning can be illustrated in a thought experiment adapted from work by Wynn (2002). It relates to why incompleteness cannot justify changing constituent rules for architectures. Suppose we were to start producing Acheulean artifacts as a hobby tomorrow. Having made some, we put them in a secure case where they are found by a cognitive archaeologist in the future. This archaeologist could proceed to structure a target level where the mental architecture accounts for the entire structure of our mind, while the behavioral architecture represents the behavioral pattern underlying stone tool making. To put it bluntly, the implemented theory that results in this case by vertical and horizontal mapping could lead to the conclusion that we have a mental architecture comparable to that of the Acheulean populations. This is the result of incongruence based on the fact that the whole mental architecture (the mind) is compared with only a fragment of the behavioral architecture (tool making). However, that same future archaeologist would also most likely find the remains of roads, rocket launch sites and abstract sculptures forged out of carbon fiber. Armed with our method and a well-formed theory, the fuller set of traces of the behavioral systems that made up our material culture would lead our archeologist to infer that our mental
architecture was likely more advanced than those of Acheulean populations. Further examples of incommensurability linked with improper construction of architectures will be discussed in later sections.

A solution to incommensurability that can emerge from the use of incongruent architectures can capitalize on wider variation in a behavioral system that we noted earlier. The many and varied behavioral practices that are part of a species' repertoire can be used to repeat the mapping process and counteract the problem of incompleteness in a single behavioral architecture. The final conclusion from such a sequence of operations would be the same as a proper mapping with congruent architectures. However, in our view this process is not necessary and can be avoided by simply integrating all iterations in a single poly-architecture. A poly-architecture can be defined as the addition (Σ) of all the individual behavioral practices and systems in a synthetic architecture that accounts for the entire behavioral repertoire (Fig. 9) as well as for the entire brain circuitry (not represented).

The poly-architecture [A*] is the aggregate of the partial architectures and the same strategy is used for all the subsystems within the layers of reference of the same architecture. Each [B*] now stands for all the behavioral processes that are at the base of a single practice, while the [C*]s are the constituent elements that interact to produce those processes. This extension of Barnard’s diagram now allows us to correctly construct architectures, so that they will be congruent and the resulting horizontal mapping will be coherent, as well as the final stage of vertical mapping for the construction of an implemented theory (Fig. 5). In further developments of this methodology, the logic of such aggregations could be open to more formal representation within the mathematics of set theory. If the horizontal mapping and the construction of the meta-theories are coherent, then it means that the call for explanations from the core-level is correct as well and the mapping can then produce candidate implemented theories, that can be properly submitted to the criteria for selection for plausibility we specified earlier.

**Problem of universality**

Up to this point, our agenda has been focused on the attempt to define a universal methodology for CA at the normative level. In the next sections, we will shift to the descriptive level in order to create a concrete perception of concepts and mechanisms that have been rather abstract until now. In particular, we will compare the methods adopted in current CA theories with the methodology discussed in previous sections. Our goal is to provide examples of theories that conform to our proposal and theories that instead do not and therefore need to be adjusted.

As it will become clear from what follows, a significant proportion of theoretical outcomes in CA conform, at least in part, with our proposed methodology. Indeed, several theories contain parts that are commensurable and of comparable efficiency. Here, we will focus on Mithen’s (1996) modular hypothesis, Coolidge
& Wynn’s (2005) Enhanced Working memory and Barnard’s (2010a) Interacting Cognitive Subsystems. We will show that while their respective macro-theories of cognition differ among the three core-theories, the principles on which these architectures have been assembled are consistent. The proposals are therefore commensurable at least within the level of cognitive architecture. At the same time, the behavioral architecture depicted by these core-theories seems to be quite invariant, opening to the possibility of a proper comparison that includes more than one system. Since there is a paucity of evidence concerning the detail of neural architecture at present, it will not be included in the following discussion on commensurability. A great deal more research will be necessary to fill in the gaps in evidence, as well as to explore correspondences between our methodology and the extant theories in CA.

In 1996, Mithen advanced an account for CA based on evolutionary psychology and the modularity of mind argument (Barkow et al., 1992; Buss, 2005; Pinker, 1997; Plotkin, 1997; Fodor, 1983; Gardner, 1983). According to his model, human mind evolved from a series of isolated domains of knowledge, also referred to as “multiple intelligences”. These domains gradually became more interconnected with the increasing complexity of the genus Homo, until modern humans developed a module for meta-cognition. This had the ability to manipulate information flows between the other domains of intelligence. In this way, modern humans, but not Neanderthals, evolved a fluid cognition that enabled them to innovate and develop the wide range of Upper Paleolithic artifacts.

Coolidge & Wynn (2001, 2005, 2009) advanced a proposal, based around a quantitative aspect of Alan Baddeley’s tripartite working memory model. This architecture is composed of a central executive and two slave subsystems: the phonological loop and the visuo-spatial sketchpad (Baddeley & Hitch, 1974; Baddeley & Logie, 1999; Baddeley, 2000, 2001, 2003). Coolidge & Wynn argued that selective advantage would have accrued to Homo sapiens with an increase in working memory capacity. This would have allowed our species to perform a simultaneous integration of more complex information. In their view, indeed, a genetic mutation in brain networks at 90-50 kya may have enhanced working memory capacity and, in consequence, the development of complex tools and behaviors.

An alternative to Coolidge and Wynn’s model has been recently advanced by Philip Barnard (2010a) with his Interacting Cognitive Subsystems model (ICS). Within this perspective, evolution of the mind has been described as an additive process: cognitive architectures gradually became more complex via the addition of new reciprocally interacting subsystems. These developed as a result of both biological and embodied cultural dynamics. New subsystems depend on an iterative mechanism where inputs coming from two sources (e.g.: audition and body states feeding back changes in vocal musculatures) are mapped together in multimodal space, to establish their invariants that can bind them together (in this example, the invariants that underlie vocal output and heard speech is “phonology”). Once a repertoire of invariants differentiate from the bulk of other multimodal invariants, a new, functionally independent subsystem emerges (Barnard et al., 2007). The proposal is that an architecture of six interacting subsystems can fully explain the behavioral repertoire of apes, as well as that of our last common ancestor. Three additional modules, for vocal articulation, phonology and for propositional meaning, were added. This last addition brought into existence not only propositional meaning, but also augmented precursor multimodal capabilities to effectively yield two levels of meaning. This in turn enabled a dialogue between the two levels of meaning that support abstract thought and innovation. Barnard’s mental architecture can do more things at one and the same time than precursor architectures with fewer subsystems. A nine-interacting subsystem architecture can walk, talk, chew gum and think at the same time. Across the trajectory from six to nine subsystems, the computational power of the full architecture increased and with that the behavioral repertoires they were capable of exhibiting.
There is good reason to argue that these mental architectures are commensurable and could be properly compared. Indeed, even though the nature of the Basic interactors [B]s varies among the various models, the architectures are constructed by adopting the same logic. This suggested isomorphism between the three mental architectures is shown in Figure 10.

In Mithen’s evolutionary psychology model, the mental architecture [A] would be represented by a series of intelligences, which would act as basic interactors [B]s and would be constrained eventually by a series of evolutionary determined constituents [C]s. In Coolidge & Wynn’s proposal, conversely, the basic interactors would be covered by the subsystems in Baddeley’s model (visuospatial sketchpad, central executive, phonological loop), composed themselves by sub-components (for the phonological loop: articulatory rehearsal and phonological store, perhaps), while in Barnard’s model the mental architecture would be represented by the interacting subsystems (the [B]s), which are composed themselves by a number of constituent subcomponents (their internal structure), which stand for the [C]s.

If we consider, as we argued earlier, that an invariant behavioral architecture can be constructed for these three different proposals, we have the possibility to map horizontally at least two systems (i.e.: the mental and the behavioral) in a coherent way. As a consequence, we have three different core-theories, which, albeit only partial because of the absence of the neural architecture, can be used to produce properly comparable implemented theories. The efficiency of these resulting implemented theories can be judged in a two stage process. First, in the construction of the core-theories, attention would be focused on internal congruence of the architectures, conformity to the experimental data, explanatory power during the horizontal mapping, etc... Second, during the vertical mapping for the target-level, the resulting implemented theories would be evaluated for their plausibility. In this way, the number of ad hoc hypotheses that need to be produced in order to support the explanations of the properties of the two missing systems (Fig. 5) will be counted. In summary, these three commensurable theories can be compared and analyzed through the criteria for selection that we previously discussed. Thus, theories providing less efficient explanations could be ultimately rejected.

Unfortunately, apart from the small number of implemented theories reported above, many proposals in CA do not fit with our methodology. On the contrary, most of them appear to be largely incommensurable with those described above. This general incommensurability has been noted, for instance, by Wynn & Coolidge (2011), who highlighted the fact that ambiguous terms like “complex” or “modern” cognition are widely adopted nowadays by archaeologists interested in the evolution of mind and behavior. Indeed, these notions are often used in place of precise descriptions of the cognitive processes and mental architectures necessary to produce artifacts (see also Dubreuil, 2011). For example, the elaborated
sequence of stages that are necessary for performing a Levallois reduction could be considered as a proof of underlying “complex” cognition. However, this tells us little, for instance, about the working memory capacity that is needed to perform the same tool-making task. As a consequence, this difference in analytic categories clearly contributes to confusion within debates.

Problems like these stem from the fact that many theories in CA focus only on the properties of the artifactual record. Thus, properties of a behavioral architecture, or even isolated fragments of it, are used to directly draw inferences about an ill-defined structure of the mind (e.g.: complex mind, modern mind, etc...). No micro-theories of the subcomponents of the mind itself and the macro-theory that rules their reciprocal interactions are taken into consideration. In our view, this leads to a series of situations like the first type of incommensurability, where logical connections and mappings are produced with a theory-dependent logic. For instance, Barnard (2010b) reports that his diagram has a high risk of being misused through what he defines as “diagonal connections” between components of the architectures, as opposed to proper horizontal mapping. He discusses a case where specific properties of cognitive subsystems are inferred directly from isolated components of behavioral architectures. As an example of improper diagonal connections, he uses the attempt to infer properties of language from regularities in stone knapping procedures (Holloway, 1969), which is addressed without any horizontal reference to a theory of cognition.

Particular emphasis in discussing problems of incommensurability should be given to the spread use of the concept of “behavioral modernity” in the archaeological debate, an expression that itself comes with a high degree of ambiguity: there is no solid theory to account for it (Henshilwood & Marean, 2003). The main risk with attributing the label “modern” to a behavioral architecture is that it can lead to the automatic transfer of this qualitative attribution to the mental and neural architectures as well. Again, this can be performed without taking the exact nature of neural and cognitive architectures required to support the behavioral repertoire itself into consideration. For instance, defining the behavior that can be identified only in anatomically modern humans as modern tells us little about the “modernity” of the mental architecture. Even if a particular set of behavioral practices is detected exclusively in *Homo sapiens*, in fact, the cognitive processes underlying them can be thoroughly consistent with a more primitive mental architecture (Klein, 2003).

Problems like these become even more relevant when we consider the relativistic drifts that in turn plague the same notion of behavioral modernity. This is strictly dependent on the authors’ beliefs and is not based on objective criteria (see Nowell, 2010 for an extensive review; Soffer, 2009, p. 45). Modern behavior has been largely associated with the use of symbols (Chase, 2003, 2006; Davidson & Noble 1989; Noble & Davidson 1991; Henshilwood & Marean, 2003; Gamble, 1999; Marean, 2007; Wadley, 2001), but it could be possible to raise the threshold of the concept of modernity to agriculture, the advent of writing systems, or even to communication through the Internet (Chase, 2003). Or perhaps we have never been modern, as suggested by Latour in his homonymous book (1993). It appears then that it is extremely difficult to classify behavior under the qualitative category of “modernity” and any assumption in this direction inevitably holds a certain degree of arbitrariness, which grows exponentially when transferred to the domain of cognition.

From these premises it follows that attempts to infer the presence of a modern mental architecture from “modern” behavior are unsafe in principle. They ought to be replaced with proper horizontal mapping, as proposed by Barnard’s extended model.

**Case Studies**

**Neanderthal symbolism**

An example of incommensurable proposals in the extant cognitive archaeological theory
might be useful to build a clear perception of the problems previously accounted for. Both Nowell (2010) and d’Errico & Stringer (2011) recently reported on the most prominent schools of thought associated with the evolution of human behavior and mind. Of the three current schools of thought recognized by these authors, we find the situation with the “cultural” school (Chase, 1999, 2003, 2006; Hovers et al., 2003, 2006; d’Errico et al., 1998; d’Errico, 2003; Conard, 2008; Kuhn & Stirner, 2007; Zilhao, 2007) particularly interesting for our discussion on incommensurability. This school is known for claiming that demographic changes in human populations might have been the main cause of variation in human behavior instead of cognitive or genetic factors. The rise in innovations is explained by appealing to the growing number of inter-individual interactions within a wider community (Shennan, 2001; Powell et al., 2009). A striking example that might be ascribed to this framework is the recent Zilhao et al.’s (2010) work on symbolic cognition in Iberian Neanderthals, associated with the findings of perforated beads in Mousterian layers. These artifacts resemble those found with modern humans in Africa and in the Near East during the Middle Stone Age (d’Errico et al., 2005, 2008, 2009; d’Errico & Vanhaeren, 2007; Bouzouggar et al., 2007; Marean et al., 2007; Bar-Yosef et al., 2009.; Vanhaeren et al., 2006) as well as in the European Upper Palaeolithic (Klein, 2008; Vanhaeren & d’Errico, 2006).

The authors discuss the implications of the use of beads, which they consider to be symbols a priori, without providing a description of the semiotic relationship between objects, signs and interpretants (Peirce, 1839-1914 in Hoopes, 1991; Rossano, 2010; Deacon, 1997), which is desirable to precisely define the structure of behavior. Then, they draw directly conclusions about the cognitive level, as demonstrated in the following passage taken from the same Zilhao et al., 2010 (emphasis added):

"The symbolic implications of body painting and of the ornamental use of pigment-stained and perforated marine shells are uncontroversial in UP and later prehistoric contexts but, as shown by the evidence from Africa, the Near East and now Iberia, both behaviors first occur in the MP/MSA. Their emergence in two continents, among two different lineages and, in the time scale of human evolution, at about the same time, is inconsistent with cognitive-genetic explanations and implies that these innovations were fulfilling a need—aiding in the personal or social identification of people—that did not exist in the preceding two million years of human evolution. Our findings therefore support models of the emergence of behavioral modernity as caused by technological progress, demographic increase, and social complexification and show that there is no biunivocal correlation between "modern" anatomy and "modern" behavior (pp. 13, 36–38)."

From the quoted text, it seems clear that the authors are using beads to argue that Neanderthals posses the ambiguously defined trait of "behavioral modernity", which in this case is identified with symbolism on the base of "growing consensus" (Marean, 2007, p. 367; see also Nowell, 2010). Behavioral modernity, warranted by the use of beads, is thus used to infer that - cognitive/genetic mechanisms are not necessary to produce those behaviors typical of fully modern humans only. On the contrary, the authors conclude that even more archaic populations can develop behavioral enhancements by relying solely on demographic and social changes. However, no information is provided on which cognitive architecture is necessary or sufficient to produce body ornaments like beads. Nor the authors clarify how this architecture is influenced by the variation of demographic and social dynamics in Neanderthal populations. Furthermore, paleoneurological evidence (Bruner, 2003, 2004, 2008, 2010, 2011; Gunz, 2010), which would argue in favor of a cognitive and biological hypothesis, is also not taken into consideration. This is even more evident through the absence of any reference to a neural architecture.
In the light of our methodology, it can be argued that the logic adopted by Zilhao et al. (2010) is based on establishing connections between a part of a behavioral architecture, an isolated B (use of beads in Neanderthals), directly to an improperly defined structure of the mind (A = symbolic). In Figure 11, indeed, we can now visualize how a partial behavioral architecture is mapped on the whole structure of the mind, without taking into account any macro-theory or micro-theory for a mental architecture (Incommensurability of the first type).

In addition, even if mapped with a mental and a neural architecture, the logic is still incongruent in constructing the architectures (incommensurability of the second type) (Figure 12). As discussed before, single behavioral practices, considered in isolation from the entire behavioral repertoire, cannot be used alone to explain properties of a cognitive architecture.

The inappropriate mapping adopted by Zilhao et al. (2010) leads to the claim that a qualitatively modern cognitive architecture might have been present in Neanderthals as well (Harrold et al., 2009, p. 290) and discovered or exploited through changing demographics/cultural phenomena. This implemented theory radically contradicts the general conclusion shared by Mithen’s, Coolidge & Wynn’s and Barnard’s models, despite being incommensurable with them. These three proposals, albeit with consistent theory-specific differences, share the idea that cognitive/biological differences were in play among modern humans and the non modern populations.

The crucial point to grasp here is that the cultural school’s conclusions could be contradicting the alternative proposals only by virtue of the improper mapping adopted and not because they represent more plausible explanations. Once a proper mapping is adopted, the new constraints offered by the many macro and micro-theoretical levels introduced can undermine the stability of the same theory. In what follows, we will show how properly built core-theories could in principle explain the use of beads by archaic populations without supporting the cultural school thesis of cognitive equivalence.

Indeed, we can adopt Barnard’s extended diagram to construct a proper behavioral architecture, which includes the use of beads as a basic interactor, along with all the repertoire of practices and the constituent interactors that constrain them. Then, if we map this onto a mental architecture and its subsystems, as previously described, and then on a neural architecture, we can realize that the cultural school conclusions does not follow necessarily from the use of beads. In fact, all the processes that are required to produce the Neanderthal behavioral architecture (poly-architecture), including those that underlie the use of beads, could be supported by a non-enhanced working memory, an architecture of eight interacting cognitive subsystems or a set of intelligences lacking metacognition (Coolidge &
Wynn, 2004; Wynn & Coolidge, 2004; Barnard, 2010a; Mithen, 1996). Contrary to the common archaeological notion that symbolism is connected with beads, this behavioral practice could be, in effect, reduced to a non-symbolic level following a Saussarian framework (Gärdenfors, 2011), an enactivist perspective (Malafouris, 2007, 2008b), or more precisely to an indexical one, by adopting a Peircean semiotics approach (Rossano, 2010, 2011). In this way, cognitive processes at the base of indexical reference can be identified, mapped on the mental architecture and evaluated together with the rest of the behavioral practices associated with Neanderthals.

By virtue of the universal laws that bind the architectures in the core-theory, the vertical mapping then allows us to reconstruct the missing architectures by presenting an explanatory set of theories in the target level, which ends up in the formation of an implemented theory (Fig. 5). In this case, if the use of beads does not exceed those cognitive processes that can explain the rest of the behavioral architecture too (see Coolidge & Wynn, reply to Henshilwood & Dubreuil, 2011), there is no need to argue in favor of a modern mental architecture in Neanderthals.

However, the idea that demographic changes can lead to behavioral advancements, as hypothesized by Zilhao et al. (2010), can still be supported, if behavioral variations are performed within the limits of the capabilities of a non-fully-modern mental architecture. This, however, does not rule out the fact that biological alterations might have been necessary to reach a fully modern mental architecture. In addition, it does not prove either that demographic changes alone are responsible for the “discovery” of cognitive capabilities in non-modern human species.

Nevertheless, we do not intend to say that the approach of the cultural school cannot be followed or does not deserve any consideration *a priori*. Our goal was simply to compare this school of thought with the other models described above, in order to discuss an example of how theories can be incommensurable and contradictory at the same time. Moreover, we intended to point out the fact that some theories, once contextualized in a proper mapping framework, can lose their original explanatory power.

Incidentally, it appears that all the models we have argued to be commensurable in our methodology actually support the same final
conclusion. However, this does not imply that our methodology is necessarily linked to the sole cognitive/biological explanatory models. This might be an additional reason for the followers of the “cultural” school to adjust their proposals in a way that is commensurable with the other models. For example, they would need to specify more clearly: a) what happens to mental and neural architectures when the behavioral one varies in response to demographic and social changes, b) which constraints prevent a modern cognitive architecture from being exploited in the absence of demographic changes, c) how current macro-theories of cognition explain the discovery of latent cognitive processes, d) what the difference in parietal lobe anatomy between modern humans and Neanderthals means (Bruner, 2004, 2008, 2010) and so forth. On the other side, the cultural school can contribute to our methodology by providing a framework to define an assembled system of social and demographic mechanisms/theories (Shennan, 2001; Powell et al., 2009). This could exert further constraints on the behavioral architecture and on the whole process of generating core-theories.

Based on this approach, it would be possible to create candidate implemented theories, which might be properly compared with alternative theoretical proposals and evaluated for plausibility in light of our criteria for selection. In the current state of affairs, the fact that the cultural school offers theories that are mostly incommensurable with those from the other schools of thought demonstrates that methodological relativism represents the status quo in cognitive archeology. This example embodies how deeply distorted (and inefficient) a theoretical debate based on contradictory and yet incommensurable theories is.

Use of beads in early modern humans

Recently, Henshilwood & Dubreuil (2011) advanced a proposal on a related topic, namely the use of beads in Early Modern human populations in South-African Still Bay and Howieson’s Poort technocomplexes. Interestingly, their approach to the potential symbolic implications of this practice can be considered as broadly consistent with our methodology.

The authors provide a multidisciplinary analysis of the controversial problem of the use of beads in Early-modern Homo sapiens populations. They describe a behavioral architecture that includes the production of beads, advance explanations on the cognitive processes that are compatible with the use of beads as ornaments, and propose a neurological explanation for these cognitive requirements. Henshilwood & Dubreuil argue that beads can be considered as symbolic (p. 375) and that this explanation is supported by a series of cognitive processes that implies the presence, within these populations, of a working memory capacity that almost equals that of contemporary humans (p. 379). This is furthermore supported by other synchronic archaeological evidence (pp. 372-375 and references therein; Wadley, 2010). This hypothesis is also corroborated by their argument that there was a rise in complexity of the temporal lobes, which is in turn supported by studies in evolutionary neuroscience that show a disproportional enlargement of the temporal volumes in modern humans as opposed to apes (Rilling & Seligman, 2002; Semendeferi & Damasio, 2000). In addition to volumetric variations, they also refer to the enhanced functional characterization of temporal cortices, their pattern of connectivity and development (pp. 362-367 and references therein).

Comparing Henshilwood & Dubreuil’s proposal with that of Coolidge & Wynn (2011), it can be argued that these authors refer, more or less directly, to the same set of theories in constructing a core-theory (the same set of rules), which might be associated to Baddeley’s framework (Baddeley & Hitch, 1974; Baddeley & Loogie, 1999; Baddeley, 2001, 2003). Even though this does not exactly match our proposal for producing assembled systems, their macro-theories and the mapping processes, it is important to recognize that we agree in the methodological direction that ought to be taken. It is worth noting that every assertion these authors make with respect to behavior is connected to explanations in terms of executive functions and cognition, as
well as neural substrates. This creates a series of inferences that work as a holistic network.

However, these groups of authors interpret the behavioral architecture in different ways, so that the relationships among the entities within the behavioral architecture are different (one allows symbolic exchange of information between entities, the other only indexical). As a consequence, the result of vertical mapping, along with the explanations drawn from the core-theory, both differ. In this case, the same core-theory produces different conclusions depending on how the behavioral architecture is constructed from the artifactual record. The result of this vertical mapping is the formulation of two candidate implemented theories, which are perfectly congruent in the rules of construction of the architectures, internally coherent in the mapping processes and ultimately commensurable when compared. These two implemented theories, albeit contradictory in their conclusions, can then be submitted to our criteria of selection for plausibility.

Which conception of the behavioral architecture best fits with the artifactual record? How many ad hoc hypotheses are requested to make the implemented theories stable? Do they leave some gaps of knowledge behind? Are they simple or is it necessary to multiply the assumptions required to support them? Questions like these can be answered now by examining the difference in the contents of the two implemented theories, not their rules of construction.

It is worth noting how the two groups of authors defend their respective models in the context of the commentary to Henshilwood & Dubreuil’s work (2011). Coolidge & Wynn claim that Henshilwood & Dubreuil’s implemented theory is implausible, because ad hoc hypotheses are required to associate beads with symbolism and modern cognition. For instance, to support the idea that beads are processed like true symbols, and not like indexes, it might be necessary to state clearly what the abstract concept embodied by these artifacts is, how this association is cognitively mediated, which cognitive processes are involved and how the symbol is inserted in a system of symbols that allows it to be considered this way (Deacon, 1997). Without providing a detailed account on these points, the association between beads and symbolism seems to be unjustified.

Henshilwood & Dubreuil defend their position by stating that the symbolic value of beads is explained by associating the abstract concept of “coolness” to these artifacts, as well as to contemporary ornaments like earrings, a notion that is nevertheless rejected by the former authors as an ad hoc assumption. Henshilwood & Dubreuil might need a further ad hoc hypothesis to clarify why a mental architecture, assumed to be capable of handling metacognitive tasks in the creation of symbols, kept this potential latent for thousands of years after reaching this level. From their side, Wynn & Coolidge have to deal with problems traditionally connected with every mutation-based proposal. For example, they need to provide additional clarifications on the chronology of mutational events, the speed of replacement of alleles that specify for non-enhanced working memory phenotypes, the outcome of behavioral advantages, etc... (see for a review Wynn & Coolidge, 2011).

Conclusions

Our lack of access to the mind of ancient populations, as well as the necessity of reconstructing its nature by mapping together multiple and reciprocally interacting systems, represent probably the most important epistemological issue in CA.

Scholars who have chosen to embrace a non-reductionist view and to take these problems into consideration might, therefore, be tempted to overcome them by accepting methodological relativism. In this paper, however, we have argued that the intrinsic limits of this discipline do not require scholars to adopt a subjective/interpretative approach. As an alternative to relativism, we have proposed a method based upon a synthesis of the following:

1) **Horizontal Mapping - Meta-Theoretical Level:** The construction of universal laws of...
connection between the interacting systems examined in CA (in our case neural, mental and behavioral architectures) based upon empirical and theoretical research in extant species. These rules should allow the prediction of how any potential set of systems interacts, so that, given one system, and constraints that govern their connections, the nature of the other two can be inferred.

2) **Construction of the Target Level**: The individuation and definition of a behavioral architecture, its own constitutive elements and the way they interact. This operation is mediated by the analysis of the archaeological record associated with the target species/population of interest.

3) **Vertical Mapping - Implemented-Theory**: The deduction of missing architectures and their interactions in the target level. This is performed via universal principles of connection determined by point 1 above, as applied to the target level’s behavioral architecture (point 2). As a result of this procedure, we have what we have defined an “implemented theory”. This is not necessarily valid in the explanations it provides, because elements of implausibility can arise out of each stage of construction (point 1 and 2). However, this resulting implemented theory is logically coherent and can represent a candidate theory to be compared with other candidate proposals in terms of plausibility.

4) **Theory selection**: Plausibility of candidate implemented theories is evaluated on the basis of criteria for selection and the most plausible theories selected given our state of knowledge.

This methodological proposal demonstrates that, even in the absence of strict empirical reference, it is still possible to adopt a methodological framework that allows us to evaluate theoretical proposals on the base of rationally objective criteria, which markedly differ from subjective interpretations.

In summary, within the wider field of CA remains a risk of reintroducing relativism, which would lead to the dangerous paradox of “formalizing” the idea that anything goes. To counter this risk, we have proposed an extension of Barnard’s (2010b) model. Our proposal provides a precise logic to be followed when mapping domains of knowledge in CA. This method is still incomplete and may include additional assembled systems. Nevertheless, we argue that it is not only more efficient than any relativistic perspective, but it is also necessary in order to allow candidate theories to be properly compared. In this way, the risk of comparing incommensurable theories, built from different epistemological bases, can be avoided. Indeed, incommensurability can lead researchers to use a subjective methodological perspective to correct or ignore possible conceptual problems that arise within a theoretical proposal. In addition, adopting a distorted methodology can lead scholars to focus their attention and efforts on models that might be flawed in their contents. These models can be perpetuated by virtue of this methodological flaw, but can then divert effort and focus away from properly constructed models.

Of course, it is not our position that all the studies in cognitive archaeology which are not aligned with our proposal should be totally dismissed. Even works focusing on single or partial architectures can, in fact, contribute valuable new evidence, problems and questions concerning the nature of the systems of interest. These elements can be nonetheless contextualized within our model and can contribute to the formation of more inclusive and more properly constituted theories.

In addition, we are not claiming that our approach is the only one admissible as a universal methodology in this discipline. Instead, we hope to stimulate the development and the discussion of further techniques able to provide more efficient methodological tools than those discussed here. In our view two points are absolutely critical: i) providing efficient ways of establishing logical connections between/within assembled systems and ii) proposing rational criteria for evaluating commensurable theories. These elements represent the foundations of an anti-relativistic agenda and must be adhered to.
Proposals like ours are inevitably affected by the problem of dealing with real communities of scientific practice. These are usually prone to self-determination of their own methods and often unresponsive to potential suggestions of changes from outside their communities (Bell, 1998). However, our short-term goals are two-fold. First, we hope that this paper will drive the attention of cognitive archaeologists to the problems discussed here, encouraging new ideas. Second, we wish to provide a tool for identifying theories that carry elements of incommensurability and that consequently need to be further developed.

Glossary

**Architecture**: An aggregated system of structural entities that interact by virtue of their properties, creating a behavioral outcome.

**Arcuate Fasciculus**: A neural pathway that connects posterior regions of the brain (temporo-parietal junction) with the frontal cortex.

**Core-theory**: An abstract representation that defines a level where all the meta-theories are integrated together, in order to highlight the invariant rules of system interaction that are valid for all the considered species. The result is a theory that accounts for how neural, mental and behavioral systems interact and constrain each other in all potential conditions, given the evolutionary links that connect organisms.

**Horizontal mapping**: The process of aligning architectures and establishing connections between their constituent parts which follows a universal logic that is determined *a priori*.

**Implemented Theory**: The entire theoretical schema that results from the process of horizontal and vertical mapping and provides explanations about the nature of ancient minds.

**Incoherence**: In general, this term has been adopted to define a distorted mapping procedure. More specifically, in the horizontal mapping, this is the result of mapping between incongruent architectures or diagonal inferences that put layers that are not allowed to interact into communication (due to differences in the mapping rules between systems). In the vertical mapping, this indicates both the propagation of an incoherence in the horizontal mapping phase, as well as the assumption of different rules for mapping architectures between the core and the target level.

**Incommensurability**: This occurs when two or more implemented theories, each with different rules for mapping between and within systems, are built and thus cannot be compared because of the differences in the methods of construction.

**Incongruence**: A term adopted to describe using of different rules of construction for single architectures, where the internal layers are built by referring to different levels of reference, so that each hierarchy is different than the others (i.e.: different in terms of the rules of construction within each system).

**Level**: The imaginary context where architectures are aligned and mapped together by virtue of their properties.

**Macro-theory**: A theory that structures a system of local theories in a holistic whole and not the sum of reductive units.

**Meta-theory**: A set of theories that are mapped together for a single species of interest and predicts how the constituent architectures interact and constrain each other in every contingent situation that characterizes that species.

**Micro-theory**: A theory that is generally produced in experimental situations and accounts for specific subcomponents of an aggregated system/architecture.

**Post-processualism**: A school of thought in archaeological theory that denied the possibility for archaeology to reach objective conclusions by adopting a scientific method. According to the post-processualists, indeed, archaeology was limited only to a subjective/interpretative perspective.

**Processual School**: A movement in archaeological theory that supported the idea scientific method could be applied to archaeological research. Hypotheses could therefore be advanced by collecting quantitative data and tested directly from the archaeological record.
**Vertical mapping**: The process of connecting a behavioral architecture based on the artefactual record to a core-theory, in order to produce explanations on the nature of extinct minds and brains.

**Acknowledgements**

We wish to thank Philip Barnard and David Duke for their helpful comments and suggestions on the theoretical foundations of the arguments exposed in this paper. Any mistake remains our own. We thank also Philip Barnard, Iain Davidson, Krista Dotzel and Alessio Garofoli for organizational and material help during the preparation of this work.

**References**


Peikoff L. 1981. May be you are wrong. The Objectivist Forum, 2: 8-12.


Associate Editor, Emiliano Bruner