Spanish Journal of Palaeontology

Spanish Journal of Palaeontology 39 (2), 125–146, 2024
https://doi.org/10.7203/sjp.29838
Sociedad Española de Paleontología ISSN 2255-0550 / eISSN 2660-9568

OPEN ACCESS INVITED PAPER

Is palaeoanthropology being built on scientific foundations?

¿Está la paleoantropología construida sobre fundamentos científicos?

Manuel DOMÍNGUEZ-RODRIGO & Luis ALCALÁ

Abstract: The comparative method in palaeoanthropology has been predominant over the hypothesis-testing scientific method. Multiple interpretations over the same phenomena commonly co-exist given the relaxation of demarcation criteria. The contingent nature of non-reproducible phenomena in palaeoanthropology and the epistemological limits of the comparative method for addressing structural and more systematic palaeobiological processes have made it difficult to discern the extent to which palaeoanthropology is guided by scientific protocols. The imprint of corporative and neo-liberal policies in academia has found in essentially non-experimental disciplines a good field in which academic dynamics are ruled by patronizing networks of academic elites that produce and maintain trends, paradigms that do not need to be subjected to heuristics, research agendas frequently with specific political content, and control of the discourse of the past. Foucault argued that such a power-implanted structure is not intentional but emergent at the institutional level. This translates into a palaeoanthropological praxis that thrives on fact-collecting, and which produces interpretations that cannot be checked against any demarcation boundary to test their heuristics. This is reflected in the fact that paradigms may emerge and disappear in palaeoanthropology without need to confront them with their empirical content. Some examples are presented that justify this assertion.

Resumen: El método comparativo en paleoantropología ha predominado sobre el método científico de contrastación de hipótesis. Dada la relajación de los criterios de demarcación, suelen coexistir múltiples interpretaciones de un mismo fenómeno. La naturaleza contingente de los fenómenos no reproducibles en paleoantropología y los límites epistemológicos del método comparativo para abordar procesos paleobiológicos estructurales y más sistemáticos han dificultado discernir hasta qué punto la paleoantropología se guía por protocolos científicos. La impronta en el mundo académico de las políticas corporativas y neoliberales ha encontrado en disciplinas esencialmente no experimentales un buen campo en el que las dinámicas académicas están regidas por redes clientelares de élites académicas que producen y mantienen tendencias, paradigmas que no necesitan someterse a heurística, agendas de investigación frecuentemente con contenidos políticos específicos y control del discurso del pasado. Foucault argumentó que tal estructura implantada por el poder no es deliberada sino emergente a nivel institucional. Esto se traduce en una praxis paleoantropológica que se nutre de la recopilación de hechos y que produce interpretaciones que no pueden contrastarse con ninguna frontera de demarcación para poner a prueba su heurística. Esto se manifiesta en el hecho de que los paradigmas pueden surgir y desaparecer en la paleoantropología sin necesidad de confrontarlos con su contenido empírico. Se presentan algunos ejemplos que justifican esta afirmación.

Received: 21 February 2024
Accepted: 8 November 2024
Published: 2 December 2024

Corresponding author:
Luis Alcalá
alcala@parqueciencias.com

Keywords:
Methodology
Heuristics
Paradigms
Praxis

Palabras-clave:
Metodología
Heurística
Paradigmas
Praxis

INTRODUCTION

Defining the object of study

"Regardless of their differences, nearly all anthropologists are diffident regarding theory […] Actually, aside from vague "grand theories", such as evolutionism, diffusionism, functionalism, conflict theory and cultural materialism –all of them largely programmatic hypotheses that have inspired fruitful research projects– anthropology contains hardly any theories, that is, hypothetico-deductive systems. The typical research project in anthropology is a fact-finding mission […] The normal outcome of such fieldwork is a descriptive report. Most anthropologists stop here, and some claim that this is all there should be. However, many anthropologists wish to understand what they record. And this calls for a clear idea of the very nature of both explanation and anthropology".

© The Author(s) 2024. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (for details please see http://creativecommons.org/licenses/by/4.0/), which permits use, copy, adaptation, distribution, and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source.

This is how Bunge (1999) started his critique to the scientific praxis of anthropology, nested within a general critique to the so-called social sciences. Bunge emphasized that the first step for the application of a scientific method was the definition of the object of study. He underlined that "the very first problem that anthropologists face is one they share with philosophers; namely, what is a man? Or, equivalently, what is human nature? Since humans are studied by a large number of disciplines, it should not come as a surprise to learn that every one of them yields its own partial view of man" (Bunge, 1999, p. 48). And as philosophical as this sounds, defining "human" is essential if we pretend to unravel the "human" evolutionary process. Palaeoanthropologists use the term "human" on a daily basis and most commonly, they fail to define what they mean by that, other than when they refer to modern individuals of *Homo sapiens*. As far as we know, there may be a myriad of definitions in the heads of anthropologists and although substantial overlap may exist in concept, this may co-exist with wide divergences in meaning. Again, "the problem of the proper view of man –or the adequate definition of the concept of man– is not only of theoretical interest to science and philosophy. It is also practically important, for it comes up in the daily work of the palaeoanthropologist" (Bunge, 1999, p. 51). If in the case of cultural anthropologists, they can be excused for not making explicit definition of the term "human", because there is some consensus that human is everything living people do, in the case of palaeoanthropologists, this term must be clearly defined. If we are looking at behaviour, depending on what we understand by "human" we may place it as a quality in some Miocene primates or something only applicable to the late Pleistocene and onwards. This could also apply to the biological aspect of palaeoanthropology. Although researchers constantly use the term "hominin", they rarely make it explicit. A set of features formerly thought of as unique to "hominins", such as increasing dental enamel thickness, reduction of canines and their apical wear or remodelling of postcranial skeleton to accommodate bipedalism, have become more ambiguous with the discovery of some of these characteristics in certain Miocene hominoid primates. Small canines and apical wear are documented in several Miocene hominoid specimens (*e.g.*, *Sivapithecus, Ankarapithecus, Kenyapithecus, Gigantopithecus*) (Wolpoff *et al.*, 2006). Likewise, thin enamel is not only a feature of modern apes, but it is also documented in some purported hominins (*e.g., Ardipithecus*) (White *et al.*, 2009). Although probable, compelling evidence for bipedalism in early hominins is controversial at best (Haile-Selassie, 2001; Brunet *et al.*, 2002; Wolpoff *et al.*, 2002; Galik *et al.*, 2004; Gibbons, 2004; Wolpoff *et al.*, 2006; Harcourt-Smith, 2007; White *et al.*, 2009; Wood & Harrison, 2011; Almécija *et al.*, 2013). In contrast, features associated with bipedalism and hand-gripping capabilities have been documented in Miocene apes like *Oreopithecus* (Rook *et al.*, 1999; Moyà-Solà *et al.*, 1999). Although habitual bipedalism in *Oreopithecus* has been criticized (Russo & Shapiro, 2013), habitual bipedalism cannot be supported in *Ardipithecus* either, based on the lower limb morphology of this "hominin". This shows that there are big evolutionary convergences in apes, both extant and extinct. This blurs the boundary of what constitutes a hominin. Even if admitting that bipedalism could be the qualifying mark, some "hominin" taxa lack evidence thereof because of paucity of fossils (*e.g., A. kadabba, A. deyiremeda, A. garhi*). Therefore, the most common assumption about the term "hominin" is that it refers to a creature that is ancestral to humans and not to chimpanzees, and this assumption results from purported processes that are thought of as exclusive of our evolutionary lineage, without phylogenetic systematics to be able to provide a non-confounding basis. Definition of terms would not only help to specify the object of study, but it would also allow researchers to be on the same conceptual grounds.

On the nature of scientific knowledge

Methodology and epistemology. Another element that characterizes the scientific approach is that it is based on a scientific method. Most epistemological schools coincide in defining this as a deductive-based hypothesis-testing procedure, in which hypotheses must have empirical content and validation. Hypotheses systems (*i.e.*, inter-related hypothesis series) are necessary to build theories. Both theories and hypotheses must use logically-based propositions that must be consistent and empirically testable. Testing will produce validation or rejection of specific hypotheses, and theories will contain variable amounts of knowledge (*i.e.*, heuristics) according to their number of validated hypotheses. These research programs will compete to explain and forecast facts and the ones with higher heuristics should theoretically prevail over the others (Lakatos, 1978; Bunge, 1998; Niiniluoto, 1999; Popper, 2002; Niiniluoto, 2012, 2013; Bunge, 2017). The history of science tells us that this is the way science progresses in the long term. However, on smaller time scales, science is far from progressive. Kuhn (2012) stressed that paradigms (what Lakatos identified as scientific research programs) are viciously defended against evidence by their adherents (Max Planck also claimed that academic ideas only died with their defenders, and that established scientists were the main obstacle to science progress). If this is documented in natural sciences (which have a greater component of replication and experimentation) it certainly occurs within social sciences, probably at a greater scale (more below).

The idea that there is no scientific method, following Feyerabend´s anarchic view of science (Feyerabend, 1993), has gained momentum in modern "philosophers of science" by taking advantage of two fatal circumstances: the increasing dominium of postmodern (*i.e.*, hermeneutic, textual, etc.) philosophy, and the fact that most modern philosophers are neither scientists nor have any training in it. Bunge (2015) emphasized that to be a good philosopher of science, one should first be a scientist, if only to understand how basic scientific procedure and its formal logic basis operate. Following this reasoning, one could argue that many philosophers of science could be among the few experts who could make a living without any factual knowledge about their expertise. A simple but profound premise invalidates their view: without demarcation criterion (*i.e.*, definition of what is considered accepted or rejected in a scientific method) there is no way any difference between science and pseudo-science can be implemented. Without demarcation, there is no way we can measure how far any interpretation is from reality. Without demarcation, there are no rules to interpretation and, therefore, imagination can come rampant making it impossible to differentiate reality from fantasy (some confound Bunge´s example of axiomatic theory on the existence of ghosts to argue that anything can be framed scientifically to support the lack of a scientific method –*e.g.*, Diéguez (1998, 2020)–, when what Bunge did was to use this example just to show the opposite: not all methods are valid and not all can be labelled as scientific; demarcation is essential. Several postmodern philosophers solve the science/pseudo-science paradox by saying that instead of one scientific method, there is a plurality of methods depending on the discipline. This only atomizes the problem, by making it necessary to define multiple demarcation criteria in each field. It also presents the contradiction that what is considered "scientific" in one field may be considered "non-scientific" in another, making demarcation also relative and absurd. Knowledge is absolute, regardless of what field we are interested in knowing. How we gain knowledge should not depend on *what* we enquire, but *how* we enquire it. None of the philosophers sharing this atomistic postmodern view have been able to provide a single example of a specific scientific method for a specific discipline. They gorge on theory, but science is all about praxis.

The simple fact that scientists gain knowledge when applying some protocols, but not when applying others is a devastating example to postmodern fundaments because it shows that: a) no everything is valid (against Feyerabend), b) some methods produce new knowledge whereas others produce stagnation and, c) those methods that produce knowledge do so in multiple disciplines (*i.e.*, they are not field specific). When observed in perspective, there is no doubt that science has progressed century by century (starting in the XVIIth century). However, Popper´s view on the gradualistic progression of knowledge was naive in view on how scientists operate defending paradigms (Kuhn, 2012). More than an oblique line (comparing time with knowledge progression), science looks more like a staircase with steps whose width depends on how strongly paradigms have resisted the passing of time. Here, most postmodern philosophers fail to view that science is composed of two different interactive dynamics: knowledge seeking and scientific community dynamics. The former depends on method, and is the active object of study of epistemology. The latter is social, political, and can be understood in the framework of complex systems, mass psychology and sociology (including history of science, very frequently mistaken with philosophy of science). In any academic social network, where relations of agents are based on politics, power dynamics are more influential on the development of events than methods. It is here that postmodern philosophers get most of their "heuristics". It is also here that Kuhn gets support for his view of how scientists operate and how science progresses. It is this that conditions how wide the steps in the progression of knowledge are as time goes by. Scientific communities impose the tempo at which we gain knowledge. The scientific method determines how we gain knowledge.

A recent study shows that in non-experimental and low-power sciences, true theories may not be adopted because false paradigms persist due to the homophilic behaviour (*i.e.*, favoring ideas and people that support one´s paradigm) of their practitioners (Akerlof & Michaillat, 2018). This is especially prominent in social sciences. In such cases, "science moves away from the truth" (Akerlof & Michaillat, 2018, p. 13228). This study shows that low-power disciplines do not only slow down scientific progress, but they create their own dynamics resulting in false paradigm entrenchment and scientific stagnation. Examples from anthropology that fit this situation abound. This is also documented in some "hard" sciences, such as physics: "Despite the absence of experimental support and precise formulation, the theory is believed by some of its adherents with a certainty that seems emotional rather than rational" (Smolin, 2007).

Another recent study showed "(i) that natural scientists tended to express more strongly realist views than social scientists, (ii) that history and philosophy of science scholars tended to express more antirealist views than natural scientists and, (iii) that van Fraassen's characterization of scientific realism failed to cluster with more standard characterizations" (Beebe & Dellsén, 2020). This stresses the profound methodological and conceptual divide between: a) natural-experimental sciences and the other disciplines, b) those who conduct scientific praxis and those who theorize about it and, c) the mischaracterization of realism by its critics.

Despite the many epistemological schools that emerged during the XXth century, three were the main influential think-tanks for defining scientific methods and their praxis: logical positivism, logical empiricism and scientific realism. Logical positivism was predominant during the first third of the XXth century. It based its demarcation premises on the reductionist method of considering only empirical facts and logical propositions (analytical, mathematical) as scientific.

They introduced the “meaningfulness” criterion. A proposition was empirically meaningful if it was either confirmed or presented degrees of confirmation, which others referred to as being empirically verifiable. This called for a method based on empirical contrasting. Such a method aimed at being universal and unified across disciplines. Logical positivism rejected metaphysics and causality and was initially founded on induction. The great contribution of logical empiricism and positivism was: a) the intent to unite disciplines through a single method, b) the definition of propositions following the language of logic and, c) the necessity of contrasting empirically propositions. It reduced everything that is knowable to the realm of experience.

However, this epistemological school evolved from the self-awareness of the limitations of its main precepts and the realization that the main fundamental propositions were either incomplete or incorrect (this implied the demise of the verifiability criterion of meaningfulness, the differentiation between theoretical and observational, and the analytic/synthetic distinction). This is how it morphed into Hempel´s logical empiricism, which was based on the hypothetico-deductive method and rescued logical metaphysics. According to logical positivism, scientific theories are reduced to conceptual constructions of observable events. For Hempel, this was incorrect. Hempel was mainly concerned with how we elaborate scientific theories. Theories exist and are needed to account for empirical phenomena. In essence, theories are metaphysical constructs. Theories contain internal principles (they are logical and non-empirical) and bridge principles which link them to the empirical realm. The epistemological foundation of logical empiricism lies in the way it defines explanation as scientific. Hempel based it on the “nomic expectability” thesis, which consists of deriving explanation of singular events (*explanandum*) from lawlike sentences tied to premises containing initial conditions (*explanans*). In scientific explanations, assuming that all the sentences in the *explanans* and *explanandum* are true: a) the *explanandum* must be deductively derived from the *explanans*, b) the *explanans* must be based on at least one general law and, c) the *explanans* and *explanandum* must contain testable empirical content. This deductive-nomological explanation was complemented with the “inductive-statistical” approach to probabilistic inference. For Hempel, scientific explanation was performed within the covering-law framework. No scientific explanations exist without universals. Hempel also emphasized that without demarcation, science does not exist.

Scientific realism (**SR**) expands these concepts and adds a new dimension to them. Reality adopts an observable and unobservable nature. Scientific theories, thus, yield knowledge about the universe, including unobservable parts of it. SR is built upon the premise that mind-independent reality exists (contra idealism), that such a reality can be approached and understood, and that scientific theories are the best way to provide descriptions of it. The multiple flavours of SR also emphasize the use of hypothetico-deductive methods for hypothesis and theory elaboration (Popper, 2002; Niiniluoto, 1999, 2012), but do not undermine the important role that induction plays as inspiration for hypotheses, which are deductive systems (Bunge, 1998, 2015, 2017). SR argues that some scientific theories are more successful than others because they are better approximations to the truth and can, therefore, make more successful predictions of events. They also contain a greater load of heuristically-explained phenomena (Lakatos, 1978). There are different views about whether demarcation consists of theories being scientific when they can be falsified (Popper, 2002), or corroborated, or likely or neither falsified nor corroborated but containing a higher heuristic than competitive explanations (Lakatos, 1978); however, all of them coincide in determining that scientific theories must be logically consistent, empirically testable, and compete for explaining the wider range of phenomena possible and for making new predictions. Likewise, hypothesis testing is based on the use of logically-linked true premises, which expose the conditions of the *explanans* and its path to the *explanandum*, in pretty much a similar fashion to the deductive-nomological system devised by Hempel (Bunge, 1998). Like logical empiricism, SR relies on a tight link between logically-formulated propositions and their empirical contents and testing, but it goes beyond logical empiricism by extending the implications of such empirical testing from factual hypotheses to the founder hypotheses to which they are linked through a systemic concept of theory building (Bunge, 1998). Scientific realism best explains how the physics and chemistry (and most experimental disciplines) developed in the XIXth and XXth centuries. Most of the advances and discoveries in these fields were preceded by major theoretical innovations. The atom was conceived before it was seen, and so was its functioning and composition. Einstein´s relativity theory, quantum physics, Feynman´s quantum electrodynamics are other examples. Empirical support for these discoveries came later. Still today, physical theories (*i.e.*, string theory, multiverse theory, etc.) are far ahead of what has been empirically tested and supported (Smolin, 2007, 2013; Penrose, 2017). As Hacking puts it: “Experimental work provides the strongest evidence for scientific realism. This is not because we test hypotheses about entities. It is because entities that in principle cannot be ‘observed’ are regularly manipulated to produce new phenomena and to investigate other aspects of nature.” (Hacking, 1982).

It could be argued that the three major epistemological movements share in common in their definition of scientific theory and procedure the following concepts: demarcation is fundamental to define a scientific method, the scientific method is absolute and universal (*i.e.*, non-relative according to research field), such demarcation must be built upon hypotheses that are logically consistent and nomothetic, and whose premises involve a high degree of testable empirical content. Tested hypotheses may be likely or corroborated and should ideally provide both new knowledge and predictions of new knowledge. It could be said that these epistemological schools argue that science aims at discovering generalizations, processes that operate with rules, with the secondary goal of explaining the particular with reference to the general.

These three epistemological schools are focused on the scientific methodology, that is, the procedure of how knowledge is obtained. They are not entertaining any aspiration in explaining how the human factor impacts the knowledge-acquisition process. This is the area of social psychology, sociology of science, or philosophical schools like historicism, and social constructivism. They are not incompatible if one knows that both target different objects of study. The former targets science (*i.e.*, individual praxis) and the latter targets how science is implemented (*i.e.*, collective praxis). The former depends on the scientist´s rational procedure and the latter depends on a large array of intertwined factors, which include power, prestige, politics and discourse of scientists within their communities. Ideally, scientists should collaborate to reach knowledge. In practice, this aspiration is confounded by the limitation of and competition for resources, the psychology of paradigm defence, the social group psychology, the highly-hierarchized system in which science is practised (producing a power chain), the neo-liberal academic policies in which funding, positions, promotion and prestige depends highly on productivity (regardless of quality), the confounding of scientific quality with impact-factor, instead of with production of new knowledge, the mass-media projection that science has nowadays, which results in priming social visibility over relevance, the ethnocentric and gender discriminatory dynamics of scientific promotion, and the relationship of the scientific establishment with money, politics and self-justification. This latter element has been argued to be essential for the social projection of scientists and their existential justification as an additional social class (López-Corredoira, 2013). This leads to the rather ample (and ambiguous) use of labelling somebody as a "scientist" regardless of whether he/she is focusing on producing irrelevant (or even fake) knowledge with the goal of disseminating it on the media, or somebody anonymously discovering new knowledge in the solitude of a lab. This situation also leads to the opposite of what it theoretically should contribute: scientific teams compete, discredit and fight more commonly than cooperate.

The fellowship of the ring: when knowledge is molded by academic networks

This multivariate situation in which science is practised strengthens the idea that there is indeed a scientific method, because while some groups seem to be unable to produce new knowledge, others discover new ways to explain reality. This cannot be attributed to randomness. Despite their deficiencies to explain how factual science is produced, postmodern schools of thought are well positioned to explain how these social aspects of science operate, because they are an essential part of how scientists as a community carry out science and how knowledge progresses or stagnates.

In the light of epistemology and, as importantly, the knowledge generation process in natural/experimental sciences, one pertinent question is to what degree is palaeoanthropology a science. Here, palaeoanthropology will be defined as the study of the evolutionary history of humans, both physically and behaviourally. Recently, Smith and Wood (2017) posited that most palaeoanthropological questions (especially those aiming at explaining "why") relate to contingent historical events and are outside the scientific realm. If true, this creates an interesting set of questions: if such knowledge is not scientifically tested, how is it obtained and how valid is it? If no demarcation assists in determining its heuristic power, how can it be maintained by academic communities over decades? What does this show about academic dynamics?

For the specific case of palaeoanthropology (and palaeontology in general), it has already been said that this field is a "vast superstructure built on unexamined and perhaps weak conceptual foundations" (Rudwick, 1972). This argument underscores the lack of guiding laws and hypotheses in palaeontological research, built upon a purely inductive approach, which assumes that a mere accumulation of facts will automatically generate theoretical insights, as though the theory were inherently embedded within the data itself (Gould, 1980). This has enabled the colonization of power-controlling academic elites that design the theory and praxis of the discipline. Some minoritarian voices have raised against this routinary procedure in palaeontological research (*e.g.*, De Renzi, 2005). These voices emphasize the use of the scientific method in palaeontology and the need to abandon inductivism; they require the adoption of the principle of uniformity solely from a purely methodological point of view – the timelessness of the laws of nature. The uniformity of contingent aspects (substantive uniformity) should be treated as a hypothesis to be tested (De Renzi, 2005). We join here those voices.

In palaeoanthropology, as in the broader field of palaeontology, direct experimentation is not possible in the same manner as in disciplines like genetics or physiology. Instead, hypotheses are formulated suggesting that certain events occurred under specific, well-defined conditions. These conditions are then replicated through simulations, allowing researchers to conduct experiments that test whether the events could have unfolded as evidenced by the fossil record. Such approaches are better understood as physical simulations. The formulation of such hypotheses, with a strong empirical testing approach is crucial for a scientific endeavour within palaeontology and should guide palaeoanthropological theory even from the beginning of hypothesis formulation through the design of its experimental testing. This approach, which we advocate, is still minoritarian and has been overshadowed by other factors that have more to do with the sociology of science than with science itself (see below).

BRINGING PALAEOANTHROPOLOGY IN FRONT OF THE EPISTEMOLOGICAL MIRROR

Palaeobiological aspects

The problem of classification. Most palaeoanthropological research is not structured around hypothesis-testing. Palaeoanthropologists go to the field to find fossils. They are fact collectors. Most of them do not produce a hypothesis and conduct fieldwork to test it, but collect information and then construct a way to make it intelligible. In the first phase of palaeoanthropology (not a specific time frame can be provided because it varies widely across geography), fossils were simply described. In the second phase, they were also measured and quantified. This conveyed the false idea that everything that is quantified must be scientific. This quantification was (still is) mostly carried out without any intention of testing models, hypotheses or ideas. In this regard, palaeoanthropologists proceed like unsupervised machine learning algorithms. They collect information and iterate over it to produce classification. This is how one new species of hominin is differentiated from others. Metric-analyses, conducted within described hypothesis-testing frameworks are still a minority (*e.g.*, Rodríguez *et al.*, 2010). Still, such measurements are limited in their multivariation. Only one or two (in some cases even just a few) variables are used at a time. Quantification still plays a secondary role on how new species are defined. Description is still the main guideline. Descriptive features of new fossils are then compared to other hominin fossils, but in most cases, such comparison is rather limited. Only selected hominins or fossils are used, and sometimes this selection may or may not be justified. This yields the paradox that over the same criteria, appreciation of how much any given character is present (or its extent) can be highly subjective. This leads to different palaeoanthropologists failing to appreciate the same properties when inspecting the same fossil. However, this situation underscores that comparative analysis is the mainstream way of providing interpretation in palaeoanthropology (Nunn, 2011). However, this approach fails to comply with basic demarcation requirements: how much variation is needed to justify the definition of a new species (even if using geochronological taxa)? Is the concept of species even clear in extant apes? Can we define what a species is? Without a proper definition, there is no scientific viable use of the term.

The problem of behavioural and functional analyses. Regarding behavioural inferences from hominin anatomies, the procedure is essentially similar. Evolutionary anthropologists infer adaptation patterns and behaviours from anatomical morphology and associated features. For this purpose, researchers compare hominin anatomical characteristics to those of extant primates and derive interpretations by similarity of form (and inferred similar function). The comparative method is again determinant, even when coated with quantitative phylogenetic approach (Harvey & Pagel, 1991). However, the premises used in this comparative exercise are exempt of tested (let alone general) validity. For example, how do we know that hominin curved hand phalanges imply an arboreal adaptation only because they resemble the curved phalanges of tree-climbing extant primates? This point was made by some researchers, with a lack of time-projection perspective on residual and adaptive morphology. Recent research on a primate that barely engaged in arboreal activities while growing showed indistinguishable morphology of phalanges from peers that conducted sustained tree-climbing while growing (Wallace *et al.*, 2020). This presents a currently non-soluble dichotomy: are morphological features documented in hominins a proof of their adaptation or the inheritance of primitive traits that had no relevance in their adaptation? Can we differentiate between adaptations and “spandrels”? (Gould & Lewontin, 1979). Until this question is solved, palaeoanthropologists will be limited in their interpretations of form, function and behaviour, and no scientific explanation thereof is currently possible.

Only experimentation could contribute to determining which bones and under what circumstances are epigenetically modified upon stress loading. Only experimentation could validate which specific loadings produce specific bone shapes. Wolff´s law (bones remodel according to mechanical loading) has been extensively used to interpret differences in long bone cross-section geometry and trabecular architecture as reflecting biomechanical function. However, experimental research on bone remodelling processes has shown that remodelling is highly variable in skeletal locus and that no direct relationship can be established between orientation of loads on long bones and their diaphyseal cross-section shape (Pearson & Lieberman, 2004). Exercise induce both subperiostial bone modelling and Haversian bone remodelling, sometimes simultaneously in different elements and with poor correlation with strain magnitude and locus (Lieberman *et al.*, 2003, 2004; Wallace *et al.*, 2014; Percival & Richtsmeier, 2017), in contrast with trabecular bone orientation (Pontzer *et al.*, 2006). Until now, given the lack of correlation between strain magnitude and locus loading and local modelling/remodelling processes, no relationship can be established between specific bone shapes, specific stress loadings and, therefore, specific activities. In consequence, palaeoanthropologists cannot yet scientifically infer adaptive behaviours from bone shapes. They can compare fossil to modern human populations and other extant apes, and assume (but not demonstrate) that similarities in shape indicate similarities in function; but once again, the "spandrel" problem cannot be ruled out. Palaeoanthropologists may reach a consensual state on these issues, but they cannot provide epistemological certainty on how shape reflects function. The best they can do is assume that given "the fact that we do not (and probably will not) have direct strain data available for a broad, representative array of activities, species, and skeletal locations, it is advisable to continue to use 'idealized' geometric section properties in functional analyses, with the understanding that correspondence of these with actual distributions will only be approximate" (Ruff *et al.*, 2006).

Life history, brain size and ecology. This degree of uncertainty also applies to other fields of evolutionary anthropology. For example, abundant literature on life history, brain capacity and ecology of hominins is based on analogical comparison between extant and extinct primates. A substantial amount of such inferences requires estimation of body mass. The use of molar tooth dimensions to derive regression-based estimates of body mass to be used subsequently to obtain encephalization quotiens (**EQ**) (Gingerich, 1977) is questionable because when applied to extant species it shows a great mismatch between observed and predicted body mass. The resulting observed EQ may differ substantially from the EQ derived from estimated body mass and cranial capacity (in some cases by a factor of 113%) (Smith *et al.*, 1996). Only predicted and observed body mass differ on average >30%. These differences and the regression estimates also produce wide confidence intervals that make them of limited use for application to the fossil record, since any life history feature inferred will show large variations when applying such intervals (Smith *et al.*, 1996). The differential biology of primate species, which allometrically varies from species to species, also handicaps any useful attempt to estimate body mass from fossil immature individuals (especially, infantile individuals), since it would require to justify the selection of a model as a regression reference (*i.e.*, human or ape and even justify which ape). Smith *et al.* (1996) show how the inferential chain of biological variables (from body mass to EQ, to neocortex size, to group size, to life expectancy, etc.) from one basal category (*i.e.*, molar dimensions) is biasing and not justified when using extinct primate taxa. It could be argued that, despite the strong correlation, molar area may not be not a good proxy for body mass, and alternative anatomical features might be; however, it remains to be tested in extant taxa whether the multiple proxies available to derive body mass produce similar errors or are better for inferring body mass. This is not commonly considered. Palaeoanthropologists are still engaged in deriving regressions from multiple anatomical parts and their greatest concern is if they should use a human or non-human referent when reconstructing hominin body masses (*e.g.*, Grabowski *et al.*, 2015, 2018). As stressed by Smith *et al.* (1996), this approach cannot provide confidence in how well these regressions represent body mass in extant primates and what the error is. They also produce large confidence intervals, in some cases exceeding 50 kg in range (even if using a single-taxon referent; *i.e.*, chimpanzee or human). Thus, a hominin like *Ardipithecus* may be 36 kg (19–53 kg) or 50 kg (37–70 kg) depending on whether a human or a chimpanzee referent is selected (Grabowski *et al.*, 2015, 2018), but even if justifying one (*i.e.*, chimpanzee), the probable body mass range spans almost 33 kg. Estimates of any secondary feature from such a large range is bound to produce meaningless inferences.

Phylogenetic analysis and cladistics. Phylogenetic analyses and cladistics are other important pillars of palaeoanthropological praxis. Despite its widespread use, most practitioners seem uncritical to the abundant problems that these analyses entail. Characters are commonly selected without any foundation to their biological relevance. Frequently, characters selected are inter-dependent, making the analysis redundant and biased by what in regressive statistical methods is referred to as collinearity. Characters selected can be continuous and discrete. Frequently, discrete characters are just artificial delimitations of features subjected to continuous variation. These categorizations are rarely understood within each taxa. This would be essential to understand their range of variation and true expression in one specific taxon. Gradual variation of characters is not well grasped by how characters are categorized. The foundation of cladistic analysis starts with existing phylogenies, which in extant species can be well-defended, but rarely so with extinct taxa. In this process, it is relatively easy to misidentify plesiomorphic with apomorphic characters (Curnoe, 2003). As a matter of fact, a redundant bias is the inability to separate homologies from homoplasies, which further tangles resulting phylogenetic relationships. Depending on how homologies are considered we may get widely divergent results (Ramírez, 2007). This underscores a factual problem with this method; selection and analysis of characters by different researchers may yield substantially different cladograms. The subjective factor outweighs any attempt to make the process objective.

Archaeology. The archaeological part of palaeoanthropology does not fare better. Cultural-historic approaches to diachronic variation of lithic assemblages have long assumed that typology reflect "culture" (Bordes, 1961). No anthropologist has shown this to be either a universal feature or even a valid one. Functional alternatives emerged in the 1970´s and 1980´s (Binford & Binford, 1966; Binford, 1973, 2002), which were subsequently not supported either (Beyries, 1988). More recent approaches have launched a full program to follow in the footsteps of cognitive evolution through the reconstruction of stone tool making. Cognitive analyses of lithics are almost always descriptive, including sometimes quantification of directly non-relevant features (for instance, rock mass used for knapping), based on interpretations of characteristics of flaked surfaces on cores that are strongly subjective, commonly derived from subjective interpretation of directionalities in flaking (*i.e.*, diacritic schemes), and frequently resulting in different interpretations from the analyses of the same assemblages. In this case, there is a strong experimental component, but focused on replication of artefacts, instead of the testing of hypotheses that can be objectively measured and quantified. The end product of experimentation leads cognitive researchers to support interpretations based on the authority of their experience instead of on objective data. This expert-based subjective assessment creates a situation of uncertainty when it comes to assessing different interpretations because no way to measure the heuristics of each explanation exists.

Lack of definitions of the objects of study also results in controversial academic arguments about lithic industries, which take place outside the scientific realm. A famous one is the debate on Oldowan-Acheulian in the East African early Pleistocene (Gallotti & Mussi, 2018). Some argue that any given assemblage is Oldowan, based on typology (Leakey, 1971; Domínguez-Rodrigo *et al.*, 2009; Uribelarrea *et al.*, 2017; Semaw *et al.*, 2020), and others claim that the same assemblage is Acheulian (de la Torre & Mora, 2014; Sánchez-Yustos *et al.*, 2019), based on technological criteria, which are not universally defined, but subjectively attributed to specific skills concerning flaking predetermination and planning (Tixier, 1996). The controversy is spurious since both parties are talking to each other across different conceptual universes. Most of the premises from which these discussions arise also depart from untested assumptions and have no scientific value.

One of the examples of hypothesis-testing in archaeology can be found in the application of middle-range theory in processual analysis in archaeology (Binford, 2014). The development of taphonomy in the past 40 years was caused mainly by implementing a scientific approach of referent-creation and cause-effect understanding through controlled experimentation. Although middle-range theory provided a rare opportunity to engage in hypothesis-testing praxis, archaeologists and taphonomists frequently used an incomplete concept of hypothesis synonymous with just an “idea to be tested”. This contrasts with the more elaborate epistemological concept of hypothesis, which requires, in the simplest version of scientific realism, a logically-structured set of premises (or propositions), some of which should be based on previously tested (and corroborated) knowledge, logically leading to a clearly defined consequence. The lack of such a clear definition of hypotheses leads to researchers purportedly testing the same hypotheses to end up with widely divergent results, basically because the premises (if existing) in those similar hypotheses are not the same or, more frequently, because the contextual settings where testing takes place a) do not fit the implicit premises or, b) if they do, they are different in both testing scenarios. Examples of this abound in taphonomic experimentation. For example, some studies testing meat yields at some carnivore kills suggest relative abundance of flesh in potentially scavengeable carcasses (Pobiner, 2015, 2020). Alternative experiments with the same types of carnivores and carcasses yield opposite results: meat barely survives in those scavengeable kills (Blumenschine, 1986; Domínguez-Rodrigo, 1999; Gidna *et al.*, 2014). The hypothesis to be tested is apparently the same, but the way testing is conducted differs: some experiments are done in highly-impacted anthropogenic ranches and others in national parks with far less anthropogenic impact. This impacts palaeoanthropological interpretations. For the former, scavenging for flesh would have been feasible for hominins (using that biased proxy); for the latter, that option would not have been regularly available for hominins (if the same ecological conditions are inferred; it could be argued that such inference could also constitute a biased proxy if the null hypothesis that only natural biomes minimally impacted by humans are most similar to prehistoric biomes because they show similar carcass-processing behaviours by predators regardless of their ecological diversity is proven wrong). This conveys an apparent ambiguity in the hypothesis-testing process that is not true. The situation is the result of using hypotheses that contain not clearly defined propositions and that overlook coherence between the meaning of the premises and the way they are experimentally tested (Domínguez-Rodrigo, 2012, 2015).

Bringing palaeoanthropology under epistemological scrutiny. Middle-range theory was one of the theoretical venues for applying the scientific method in palaeoanthropology. However, its praxis was (still is) inadequate. Hypothesis-testing does not commonly follow any epistemological rules. Hypotheses *per se* are not articulated beyond a main basic proposition. Meaningfulness of such proposition(s) is not evaluated and no concern for logical consistency is applied. The outcome is that what is eventually tested does not need to be the hypothesis main proposition, but something else. The original proposition is lost in translation.

In sum, most of the palaeoanthropological research is founded on the comparative method (Nunn, 2011). Information is collected, compared and then, according to parsimonious resemblance links among the compared items, interpretations are derived. This method is essential to anthropology, but it comes with serious problems. Some were summed up by Nunn (2011): a) correlation is not causation (some variables may show strong correlation between themselves, but no certainty exist if correlation is causal or accidental because of causal correlation with a third variable that is not considered in the analysis), b) comparative data is always selective (it represents samples from populations in which there is a strong input by the analyst), and most importantly, c) comparative palaeoanthropology is essentially non-experimental. This latter characteristic prevents palaeoanthropologists from conducting hypothesis-testing research. It also limits replication to the description of features and restricts objectivity (a basic epistemological requirement of science) to the minimum, if interpretations are generated and conducted with a strong subjective load.

So, back to the question of what kind of science is palaeoanthropology, it becomes easier to ascertain what the discipline is not. Hempel´s logical empiricism still contains one of the most developed concepts of what a scientific hypothesis (in contrast with a logical or formal hypotheses) should be. In order for *explanans* to account for the *explanandum*, the *explanans* must be plural and include at least one lawlike proposition. There must be a call to general lawlike knowledge (*i.e.*, a physics law). This would produce the *nomic* expectability of the *explanans*. It would also provide the certainty of corroborated (*i.e.*, empirically supported) knowledge. The construction of the hypothesis must then be based on propositions specifying how specific events or causes (C_i) can be explained by reference to general laws (L_i), upon which single events *(i.e.*, the *explanandum*) must be explained (Hempel, 1966, 2001). The *explanandum* must be deductively inferred from the set of propositions (*explanans*) and these must be true, that is, previously confirmed. As it results obvious, none of this can be applied to palaeoanthropology. The comparative method does not produce confirmed (or corroborated) knowledge. We have seen that palaeoanthropology most commonly does not base research on hypothesis testing and when it does, it does not use lawlike propositions (Hempel, 1942, noticed that “most explanations offered in history or sociology, however, fail to include an explicit statement of the general regularities they presuppose”). Therefore, no prior epistemological certainty emanates from hypotheses previously to be tested. There is also a widespread assumption that palaeoanthropology, like evolutionary biology, deals with contingency and no universal generalizations are possible. In this regard, palaeoanthropology is equated with history (Smith & Wood, 2017). This limits the range of questions to those that can be directly empirically supported, such as: what type of diet did a hominin have in view of its isotopic signal or dental microwear pattern? “Why” questions, such as the trigger of processes of speciation, emergence of new behaviours, etc. are outside any serious attempt to study human evolution. Smith and Wood (2017) provide a clear reason for this with which we agree: for those questions, palaeoanthropology misses one crucial ingredient that other disciplines have: demarcation; that is criteria to test the heuristics of explanations (*sensu* Lakatos, 1978) or capability of refuting them (*sensu* Popper, 2002).

The verdict is the same if we apply criteria from scientific realism. Even in its simplest version, scientific hypotheses should contain the following requisites for their formulation: “(i) the hypothesis must be *well-formed* (formally correct) and *meaningful* (semantically non-empty) in some scientific context, (ii) the hypothesis must be *grounded* to some extent on previous knowledge […] (iii) the hypothesis must be *empirically testable* by the *objective* procedures of science, *i.e.*, by confrontation with empirical data controlled in turn by scientific techniques and theories” (Bunge, 1998). These criteria are similar to those proposed by logical empiricism. The fact that palaeoanthropology is virtually exempt of a replicable experimental praxis basically exclude them from epistemological criteria that emphasize that scientific knowledge can only be obtained through epistemologically-correct hypothesis-testing procedures.

Is contingency the key to the historical (instead of nomological) aspects of evolutionary processes as posited by Smith and Wood (2017)? This goes against the scientific realism´s main concept that contingency aside, historical patterns show hidden processes that abide by general rules. It also goes against the logical empiricism´s concept that “history should not be concerned with the description of particular events, but with the search for general laws that may govern those events. As a characterization of the type of problem in which some historians are mainly interested, this view probably cannot be denied; as a statement of the theoretical function of general laws in scientific historical research, it is certainly unacceptable” (Hempel, 1942). According to Hempel, if history has an empirical content, then its treatment should be as in other empirical sciences. The goal is to show that events are not “a matter of chance” but that they can be explained as part of patterns and processes operating within complex systems. Under this view, palaeoanthropology is constricted by its own methodological limitations, not just by the fact that it deals with entities that no longer exist. These entities have physical and contextual properties and were, therefore, conditioned by law-abiding physical and biological processes. If palaeoanthropology could open itself to enquiry methods other than comparative frameworks, such general processes might be epistemologically accessible. One of such methods could be the use of “inductive probabilistic” hypotheses, which logical empiricism takes as scientifically adequate when no clear demarcation methods can be applied. Historical research, given its stochastic and contingent nature, may more adequately use inductive probabilistic (*i.e.,* probabilistic laws) than hypothetico-deductive (*i.e.*, universal laws) methods. A baseline of long-run frequencies of processes may be used to formulate links between events (C_i) within their probability occurrence spectra (L_i) in ways similar to deductive methods as described above.

The current status of palaeoanthropology is that it generates knowledge, but we do not have any certainty over it. There is virtually not a single aspect of palaeoanthropology where interpretations are not confronted by alternative explanations. The lack of demarcation renders them equally likely, since their heuristics are not contrasted. Given that palaeoanthropologists are capable of rejecting interpretations as they emerge or maintaining them for decades (regardless of their heuristics), what exactly determines the lifespan of interpretations and theories in palaeoanthropology? Do authority and power condition the academic dynamics of how this type of comparative knowledge is produced and maintained? Is this process related to the documented fact that in low-power non-experimental disciplines paradigms persist because of homophyly? (Akerlof & Michaillat, 2018). Why does this situation correlate positively with more anti-realistic perceptions in anthropologists compared to other natural scientists? (Beebe & Dellsén, 2020). If the quality of palaeoanthropological interpretations (and, therefore, of its research programs) cannot be objectively evaluated, then what criteria determine peer-review outcomes and funding decision processes?

A CASE STUDY: THE LOMEKWI NEW PARADIGM ABOUT EARLY STONE TOOL USE

One of the main points that we try to make in the assessment of the scientific status of palaeoanthropology, is that in the absence of objectively verifiable knowledge, in the absence of demarcation criteria, interpretations become subjective, relative, unreliable and accepted or rejected regardless of their empirical and heuristic content. In this process, academic authorities and academic network dynamics play a major role.

Current predominant neo-liberal academic agenda mandates that scientific production must be continuous. The academic factory now fulfills thousands of academic journals by providing free labour and raw material (articles) with which a corporation of publishing companies maintain a monopoly of what is scientifically published and what is not. Impact-factor pressure and produce-or-perish mantras push academic competition to boundaries in which research quality is overlooked by the same gate-keepers that are supposed to maintain it. This has led to a rate of false science higher than ever before. A study of several high-ranking journals shows that journals like Nature, Science, Cell or Lancet have between 5 to 10 times more “retracted” scientific papers than the other journals (Fang & Casadevall, 2011). It is remarkable that the “percentage of scientific articles retracted because of fraud has increased ~10-fold since 1975” (Fang *et al.*, 2012). This study analysed fraud in 56 countries and found that “The United States, Germany, Japan, and China accounted for three-quarters of retractions because of fraud or suspected fraud” (Fang *et al.*, 2012). USA alone made up about 40% of all detected fraudulent research. There is, thus, a tight correlation of scientific fraud and pressure to produce.

In early human evolution archaeology, knowledge of the big questions had remained stagnated for almost half a century. The discovery of the earliest tone tools in Ethiopia in the 1970´s supported long-held paradigmatic ideas that stone tool making and use evolved in parallel with encephalised hominins after 2.6 Ma (Semaw *et al.*, 1997). No paradigm-breaking discovery had been made in so long that researchers were ready for a change. A resurgence of the gradualistic academic trend based on the idea that the germ of stone tool use was already embedded in the brains of non-complex primates (Haslam *et al.*, 2009; Proffitt *et al.*, 2016) paved the way to the widespread acceptance of discoveries that were not substantiated. First came the discovery of purported cut marks on bones on two surface finds from Dikika (Ethiopia), but assumed to be 3.4 Ma (McPherron *et al.*, 2010). This was only marginally contested (Domínguez-Rodrigo *et al.*, 2010; Sahle *et al.*, 2017). Then came Lomekwi 3 (Kenya) (Harmand *et al.*, 2015). There, a few stone artefacts were argued to be found *in situ* in Pliocene sediments that were also indirectly dated. A tuffaceous lens in the base of a composite section (upper CSF-2012-9 outcrop, located 0.4 km south, and lower CSF-2011-2 located 0.28 km north of Lomekwi 3) is used as the fundamental layer to establish the age of the purported lithic assemblage at Lomekwi 3. The tuff is dated by geochemical correspondence with a volcanic tuff situated kilometres away. But the aforementioned key layer is not documented in the site section, neither has the composite section been stratigraphically correlated with the Lomekwi 3 one. In conclusion, as it has been already noted by Domínguez-Rodrigo and Alcalá (2016), the age of 3.3 Ma for the Lomekwi 3 stone artefacts level is still not properly supported by the geological analysis provided by Harmand *et al.* (2015). The authors use as a reference level the conglomerate with erosive base. Underlying this a tuff purportedly corresponding to the dated Toroto Tuff was identified, which itself overlies the Tulu Bor tuff. Only in one of the seven stratigraphic sections is the Toroto tuff identified. It is surprising that in two nearby sections, the Tulu Bor tuff and the conglomerate are identified, but oddly enough, the Toroto tuff does not appear in between (as expected). This is because the authors interpret a very small lense as the Toroto tuff, which apparently is mineralogically identical (they said “correlated geochemically”) to a much more widespread tuff in Koobi Fora, but it shows no stratigraphic and horizontal continuity.

Both discoveries (Dikika and Lomekwi) made it to the cover of the popular journal Nature, whose main guideline in publishing papers is their general interest and social impact. This caused immediate acceptance by the academic community and if critics existed, these were not visible. The discovery was used to turn the long tool-brain paradigm upside-down overnight. The new discovery made it to anthropology introductory books and has its own wikipedia entry. For long, only one work unveiled the absence of evidence that the artefacts were retrieved *in situ* and, therefore, that they were of Pliocene age (Domínguez-Rodrigo & Alcalá, 2016). As a response to that criticism, a subsequent publication provided more contextual information about the site and its excavation (Lewis & Harmand, 2016). With the information, what was initially scepticism, turned into convincing arguments that the purported Pliocene artefacts had not been retrieved *in situ* because the purported *in situ* items were very few and appeared clustered in the front of the eroded outcrop, within erosive gullies filled with slope deposit debris and sediment (Domínguez-Rodrigo & Alcalá, 2019). Similar observations on the derived nature of the deposition of the artefacts was also raised by Pickford (2018).

Following a classic Kuhnian example, the newly installed paradigm resisted critique. Scientific discussion should be levelled with arguments. These should consider the logical value or meaningfulness of each argument and prove it valid or wrong. One argument must be followed with another meaningful (*i.e.,* empirically-validated) counter-argument. Authoritative criteria should never be part of such argumentation. Authoritative criteria implicitly assume that one part of the debate is more intelligent than the other. Academic credentials do not validate arguments. In a detailed response to the criticisms about the "*in situ*" status of Lomekwi 3, Harmand *et al.*, (2019) only managed to defend their position with the arguments displayed in Table 1. In the evaluation of the arguments made by Domínguez-Rodrigo and Alcalá (2019), Harmand *et al.* qualified the attempt of these authors to stress the insecure basis for which Lomekwi 3 could be of Pliocene age as "professional misconduct", "spurious", "fantasized", "defying reality", "asserting with misplaced confidence", "conjectural", "baseless", "misinterpreted" "shoehorning", "imaginary", "embarrassing", "moot", "absolutely false", and "fantasy masquerading" (Harmand *et al.*, 2019). After such a scientific assessment, we expected serious empirical evidence that the critical arguments were indeed wrong. Instead, it was surprising to find that their counter-critique (*ad-hominem* attacks aside) did not come with any empirical or graphic evidence supporting any of their claims. This is especially surprising because if Domínguez-Rodrigo and Alcalá´s (2016, 2019) arguments were wrong, it would have been relatively easy to empirically show so. Their position is the weak one: they had no access to the empirical evidence of the site as the excavators do and their views are conjectural (as correctly claimed by Harmand *et al.*, 2019) precisely because of this reason. Harmand *et al.* (2019) used as evidence that the critical arguments were "spurious" the fact that they were phrased using the conditional tense ("might be …", "could potentially …", "could also …", "could be …", "would suggest …"). However, no other tense was adequate because: a) no hypothesis was presented by Domínguez-Rodrigo and Alcalá and, b) only conjectures to the published data were discussed as alternative explanations. Any other language would have been dogmatic and unscientific.

The logical scientific reaction to the critique would have been to provide empirical evidence for its rejection: a simple video or a series of photographs showing unambiguously the encased original context of artefacts before they were removed from their original position would have been decisive. These should have been, obviously, different from the rather ambiguous and controversial images that Harmand *et al.* (2015) published (see also Archer *et al.*, 2020). Additional analyses (*e.g.*, on soil micromorphology) for the geological conjectures presented by Domínguez-Rodrigo and Alcalá (2019) would also have provided support for Harmand *et al.*´s (2019) interpretations. None of this was made available. The defence to the critique was based on authoritative criteria (Tab. 1), arguments subjected to alternative interpretations and on false statements. For example, one of the latter was Harmand *et al.*´s assessment that all the artefacts they documented *in situ* referred to artefacts sealed by Pliocene sediments, which has been shown to be wrong also by Archer *et al.* (2020). Their analysis of photographs of Lomekwi 3 prior to excavation shows that several of the purported *in situ* artefacts were exposed on surface and inside erosive gullies. This supports Domínguez-Rodrigo and Alcalá´s (2016, 2019) conjectures and questions the contexts of the Lomekwi findings as well as their age. If the artefacts and photographs published so far are the best examples existing of "*in situ*" tools at Lomekwi 3, then there are reasons to question the whole discovery.

In the view of these arguments, how was Lomekwi such an instinct success? Future research may prove fruitful in discovering artefacts really *in situ* at the site. We are open to that, but at the time of writing this, such evidence does not exist and does not justify the widespread acceptance of the discovery solely by the weight of its scientific evidence. A starting point to understand the acceptance of the discovery could be the authority principle, which indicates that palaeoanthropology operates under a concept of buffered science. Interpretations have to be taken because senior figures in the field know better. Harmand *et al.* (2019) present various examples of this when they say that the critical arguments are wrong because five senior geologists say so.

Table 1. Heuristics of the arguments provided by the defense and critique of Lomekwi 3 (LOM3) as a Pliocene archaeological site.

Arguments provided by Domínguez-Rodrigo and Alcalá (2016, 2019)	Counter-arguments provided by Harmand *et al.* (2019)	Factual (empirical) evidence provided by Harmand *et al.* (2019)	Additional counter-arguments provided by the authors of this paper
Artefact C seems to be in the erosive interface between the Pliocene sediment and the slope deposit judging by the outline of the depression where it is located, which coincides with the bottom of the slope in the background	None of the sediments they annotate in red in their fig. 1 represent a slope deposit, and neither is there any 'contact' with the underlying 'light Pliocene sediment' (their description) that might be considered an unconformity […] At LOM3, slope deposit is a localized feature of only the top ~20 cm of the sub-surface, with all sediment immediately below consisting of strongly indurated Pliocene deposit.	None. A soil micro-morphological analysis could support or contradict this observation.	Red sediments in Domínguez-Rodrigo and Alcalá´s (2019) fig. 1 indicate potential erosive features (not necessarily slope deposit). Such features suggest, even to the untrained eye, discontinuity in sedimentation and discomformity. The uncomformably-overlying sediment could be of Pliocene, Pleistocene or Holocene age. No radiometric or biostratigraphic information of this unit is provided to support any age interpretation. However, whether those are uncomformable sediments of recent or Pliocene age is not relevant; that does not alter the inference that "If one projects the lowermost corner of the Pliocene sediment, it would not reach the location of C", emphasizing that its location occurs very likely at the interface of the eroded Pliocene sediment and the overlying sediment, including the slope deposit at the front of the outcrop.
This indicates another erosive structure tilting downwards in direction of the background of fig. 2D. Oddly, the core is situated in the middle of the projected rill axis (green arrow). This would suggest that the supposed *in situ* nature of the core is such, but within the erosive infill and not the Pliocene sediment. A proof thereof could be the different sediment color that can be observed between the bottom front of the fig. 2D (white Pliocene sediment indicated by A) and the darker sediment where the core is embedded (indicated by B). Compare the color of such sediment with the color in the unconformity of the 2012 wall that is not shaded (fig. 2B).	The difference in color Domínguez-Rodrigo and Alcalá identify in the upper part of the Pliocene deposit is a consequence of fine cracks and disaggregation in an otherwise massive silt, where slightly higher moisture retention results in darker color.	None. A soil micro-morphological analysis could support or contradict this observation.	The statement made by Harmand *et al.* (2919) is in contradiction with their original publication. There Harmand *et al.* (2015) refer to a large part of this darker sediment as "slope deposit" (see their fig. 2a). The authors do not indicate the criteria that leads them to take the first 20 cm of the dark sediment as slope deposit and the lower few centimeters as disaggregated Pliocene deposit. Cracks and disaggregation leading to change in color do not produce the clearly defined regular outlines that separate the uncomformable interface as observed in the images. This also explains why the outline more likely follows an erosive surface, since moisture retention does not produce ondulating sloping shapes like those documented at Lomekwi, but such shapes are typical of erosion features.
Several arguments exposed in Domínguez-Rodrigo and Alcalá (2019)	Domínguez-Rodrigo and Alcalá also incorrectly assume that sediment immediately adjacent to modern erosional features (gullies) on the land surface near the excavation consists of slope deposit. This is also incorrect: what they interpret as slope deposit is the result of the recent erosion of sterile Pliocene overburden, and is completely irrelevant to the ex or *in situ* nature of the archaeological level.	None.	In their original figures of the site prior to excavation (Harmand *et al.*, 2015), it was clearly shown that the whole area where the erosional features are documented were filled with slope deposit sediments (see fig. 2a and Extended Data fig. 5). The description of this part of the sequence as "sterile Pliocene overburden" only magnifies the possibility that the "artefacts" situated in connection with erosive interfaces may be of recent origin. Taphonomists know that erosion implies potential resedimentation. The slope deposit had rocks similar to those documented in the site. The "recent erosion" of the affected sediments could have resedimented rocks and artefacts from the overlying surface. This has major relevance to consider the *ex situ* nature of the meagre "*in situ*" assemblage. Extended Data fig. 5 also shows the extremely detritic nature of the surface sediment in connection with the "*in situ*" artefacts, which contradicts their purported occurrence and encasement in the Pliocene silty unit.
Although it was argued earlier that such unconformities were pedogenic modification of the Pliocene sediments (Harmand *et al.*, 2015), it is clear after the removal of such purported pedogenic sediments that they were infills of erosive figures. This was suggested by the clear curved outline of the unconformity, typical of erosive processes and confirmed by the gully structure of the underlying Pliocene sediment.	In this case their over- and misanalysis of photos to fit their favored geological interpretation ignores the fact –obvious to any experienced field archaeologist– that yearly rains wash over the LOM3 site to generate and compound minor surface gullying.	None.	Precisely, because the site is situated in a slope which has undergone strong erosion by the rains, the occurrence of "*in situ*" artefacts inside erosive gullies should make their possible *ex situ* nature a null hypothesis. The "*in situ*" locations of such artefacts need to be demonstrated rather than assumed. A clear example on how three of such "*in situ*" artefacts lied on the surface of such gullies, which were also covered by the detritic sediment of the slope deposit, can be seen in Archer *et al.* (2020, fig. 3)

Arguments provided by Domínguez-Rodrigo and Alcalá (2016, 2019)	Counter-arguments provided by Harmand *et al.* (2019)	Factual (empirical) evidence provided by Harmand *et al.* (2019)	Additional counter-arguments provided by the authors of this paper
Several arguments about the potential date of the slope deposit sediments.	All five of the geologists involved in the description, analysis, and publication of LOM3 to-date have independently concurred based on first-hand observation that there is no overlying sediment from the late Pliocene or Pleistocene that could be mobilized or mixed with the Pliocene tool-bearing deposits at LOM3	None. This is an authority statement, which is at odds with scientific argumentation. The arguments exposed along the next column remain unrefuted.	Similar rocks and detritic sediments to those shown in Harmand *et al.*´s (2015) extended fig. 5 in connection with the purported "*in situ*" artefacts can be observed on the slope sediments. This shows that: a) the slope sediment has indeed been mobilized and introduced inside the erosive gullies and, b) that such sediments appear indeed mixed with the purported Pliocene tool-bearing deposit. The recent analysis of the situation of the three "*in situ*" artefacts by Archer *et al.* (2020) also shows how such artefacts were on the surface of erosive gullies.
This would suggest that the supposed *in situ* nature of the core is such, but within the erosive infill and not the Pliocene sediment.	Dominguez-Rodrigo and Alcalá´s fictional 'erosive infills' on section walls have now transformed into sediments that, they now state with certainty, 'projected forward in a downsloping … manner". Both of these inferences are flat-out wrong. The absence of sediment in the center of the 'bench' results from the simple fact that in this location, Pliocene sediment was directly exposed and eroding on the original hillslope surface before excavation began in 2012.	None. Not even graphic documentation.	We never made any declaration of certainty. That would have been dogmatic and unscientific. We simply suggested alternative explanations that needed to be tested. This statement by Harmand *et al.* is contradicted by their images published in 2015 and 2016 at the beginning of their excavation, where it can be seen how the purported Pliocene sediment where the artefacts were found had already been eroded and that these rested on erosive gullies and interfaces, as we indicated and as Archer *et al.* (2020) have also shown. Those gullies can be observed in their photographs. The connections made between the gullies documented in the perimeter of the excavation and the excavated part of the site are, thus, justified.
The location of A in relation to the distant stratigraphic sections of both walls may convey the false impression that it was *in situ*, whereas this should have been confirmed by the immediate stratigraphy of the locus of the artefact, given that both sections projected forward in a downsloping (erosive) manner. Item A could also be affected by the downsloping unconformity linking the left wall in the background and the erosive feature by B on the step wall of the light Pliocene sediment.	In spite of their assumptions to the contrary, the context of the lithic piece they label 'A' was carefully determined and documented, and we reconfirm the veracity of our published data; like all excavated artefacts at LOM3, it was excavated from indurated *in situ* Pliocene sediment.	None.	That still is an unsupported statement. We assume that if such is the case, evidence must exist that shows the "*in situ*" nature of this artefact. Failure to produce such evidence leads to this interpretation not being scientifically valid because it cannot be evaluated.
Several arguments	The geological observations made from photographs by Domínguez-Rodrigo and Alcalá are incorrect, the points they raise are purely conjectural, and their argument is therefore baseless.	None.	We are open to being incorrect, but the rejection of our interpretations can only be sustained with renewed empirical evidence and not with dogmatic statements.
Interestingly, both *in situ* artefacts depicted in this image (one of them being the core shown in fig. 2D) are right in the middle of this "erosive corridor". This position, together with the inferences made from fig. 2 regarding the different types of sediments and artefact location within the erosive structure, show that every single artefact from Lomekwi 3 published to date is one way or another associated with an erosive feature affecting the Pliocene sediment.	The core shown being excavated in their fig. 2D is the exact same piece seen completely uncovered in the center of their fig. 3; seen from opposing viewpoints in the two images, it was located in the middle of the excavation floor and within the same horizon as all other archaeological materials at LOM3.	None.	We understood that also. This is why in both images we show the association of the core with a local rill (fig. 2D and also fig. 2C) which was part of a major gully (fig. 3). This reinforced our previous interpretation that the location of the piece is in the middle of the projected erosive gully. It occurs not anywhere in that gully (as shown in Domínguez-Rodrigo and Alcalá´s fig. 3) but right on the axis of it, as expected if it was resedimented from the slope deposit. This further questions that the core was *in situ*.
A proof thereof could be the different sediment color that can be observed between the bottom front of the fig. 2D (white Pliocene sediment indicated by A) and the darker sediment where the core is embedded (indicated by B).	The color differences between where the core is embedded and that observed surrounding it in fig. 2D ignores the obvious fact that sediments become dry and change color through exposure to sunlight in arid environments.	None. A soil micro-morphological analysis could support or contradict this observation.	That is true, but such a statement also obviates the fact that different sediments have different colors. As can be noticed in the Extended Data fig. 5 (Harmand *et al.*, 2015), their excavation of specific lithic items (shown in time-lapse frames) shows how the sediment color of the matrix surrounding artefacts contrasts with the substantially apparently lighter clast-free underlying Pliocene sediment.

Arguments provided by Domínguez-Rodrigo and Alcalá (2016, 2019)	Counter-arguments provided by Harmand *et al.* (2019)	Factual (empirical) evidence provided by Harmand *et al.* (2019)	Additional counter-arguments provided by the authors of this paper
Fig. 2C (shaded area) shows the direction of the erosive rill (arrow), which continues in fig. 3. This figure shows how the rill situated in the upper section of the profile shown in fig. 2C continues in a straight direction forming what could be labeled as an "erosive corridor", connecting the excavated rill in the profile to the extension of the rill away from the excavation. Interestingly, both *in situ* artefacts depicted in this image (one of them being the core shown in fig. 2D) are right in the middle of this "erosive corridor".	As seen in their fig. 3, the core is located nowhere near any ancient or modern erosive features (gullies). It was excavated from completely *in situ* sediment and was uncovered immediately after the removal of more than 2 meters of Pliocene overburden, as shown in their fig. 2C.	None.	Of course, the photograph depicting the location of the artefacts after excavation does not show that they were inside erosive features... because these had been already excavated! The projections of the erosive features surrounding the excavation suggest that the artefacts were inside such features. The core appears at a distance of about 2-3 meters from the excavation edge. Gullies and erosive features mark this edge as well as the outer perimeter of the excavation. One of such gullies can be projected from one end of the excavation to the other and the artefacts are situated right in the middle. Given that the space in between has been excavated, one cannot claim that the piece was found inside the gully. However, it certainly could be so. In the absence of any graphic support of the artefact encased by Pliocene sediment, we can only attest to what is obvious: the core appears completely excavated and resting (sediment-free) on the underlying Pliocene sediment. Its contextual information is not accessible, but certainly, by projecting the slope shape, considering the adjacent right side of the profile, the artefact could easily have been either on surface or very close to it. At that position it is very unlikely that two meters of Pliocene sediment overlied it. Two meters could be measured at the back walls of the excavation, but these do not represent the state of the sediment or its thickness in the forward projection of the slope.
The artefacts show different vertical positions suggesting they were resting on the modernly eroded surface.	We reiterate here that, without exception, all artifacts recovered to date – including all three identified as A, B, and C by Domínguez-Rodrigo and Alcalá – were recovered from a vertically restricted portion of this layer.	None.	Harmand *et al.*´s hold a broad concept of vertical restriction. In their mapped distribution, they show "*in situ*" artefacts spreading vertically over more than one meter of sequence (Harmand *et al.,* 2015; Extended Data fig. 1). Given that the original Pliocene sediments were silt, one assumes low-energy horizontal deposition. If the artefacts were originally on a silty surface, they would have been more likely to occur horizontally, unless they rested on the surface of a recently eroded Pliocene slope. The fact that they are vertically documented over more than one meter following the modern slope shape further suggests that artefacts were spreading along the vertical course of erosive gullies. This is further supported by images in Archer *et al.* (2020) (figs. 3a–3c).
Nowhere in the large excavation displayed in figs. 2 and 3 have other artefacts been found in any location unaffected by erosive processes. This could explain the rather strange void of artefacts in most of the excavated area and their occurrence in small "clusters".	This inference is absolutely false, and their wider hypothesis is utterly delusional. Their assumption that the artifacts shown in their figs. 2 and 3 are the only ones discovered during the course of excavation at LOM3 is wrong and is directly contradicted by our published data.	None.	We never assumed that the selected images were displaying the only artefacts found *in situ*. We simply argued (probably somewhat delusionally) that if the published images were the best examples of "*in situ*" artefacts, none of them showed a location of an artefact unaffected by an associated erosive gully. Indirectly, this questioned the provenience of the other "*in situ*" artefacts. We can only refer to artefacts that were singled out in the original publications. The other purported artefacts have not been published in their original position and therefore, are not available for scientific criticism.
Summary or arguments	The very idea that comparing a few photographs of an excavation is of comparable scientific merit to the first-hand observations of multiple highly experienced field geologists and archaeologists – while ignoring and misrepresenting the other 40 pages of published data pertaining to the site – smacks of the worst kind of academic arrogance.	None.	This is again an authoritative dogmatic argument. It has nothing to do with science. We produced arguments that required empirical detraction. Lacking this type of evidence, the arguments are still valid as alternative explanations to those provided by Harmand *et al.* (2015).
Summary of arguments	Domínguez-Rodrigo and Alcalá have by necessity ignored almost all of the information we have published to date concerning the geological and taphonomic context of LOM3.	None.	We have re-read Harmand *et al.*´s work and still fail to find any taphonomic information in it.

Arguments provided by Domínguez-Rodrigo and Alcalá (2016, 2019)	Counter-arguments provided by Harmand *et al.* (2019)	Factual (empirical) evidence provided by Harmand *et al.* (2019)	Additional counter-arguments provided by the authors of this paper
Summary of arguments.	Domínguez-Rodrigo and Alcalá discuss the photographs published in Harmand *et al.* (2015) and Lewis and Harmand (2016) as if these reflect a comprehensive record of artifact locations and densities, hill surface features, and the progression of the excavation. They do not.	None.	Precisely, and because they do not, they fail to support the main arguments of their interpretation. This is why our conjectures draw upon alternative explanations that question the nature of the discovery and that can be easily contested if wrong.
Age determination for the LOM3 site: A small tuffaceous lens is used as the fundamental layer to establish the age of the purported lithic assemblage at LOM3. This layer is not documented in the site section, but in nearby sections. The authors correlate this tuff geochemically with the Toroto Tuff in the Koobi Fora Formation, radiometrically dated to 3.31 ± 0.02 Ma. According to this, LOM3 would be situated 10m above the Toroto Tuff and would be slightly more recent (3.27 Ma, according to the graph in Harmand *et al.*´s fig. 3). This figure also includes a photograph in which the Toroto Tuff and the LOM3 site appear to occur on the same section. However, the stratigraphic sections of Extended Data fig. 2 lack clear correlations and section 2011-2 shows a lack of information below LOM3. The correlation among tuffs are not clear in that figure and only the Toroto Tuff is marked in one of the several sections reported (2012-9). There, it is neither correlated with the LOM3 section nor with the most complete section (2011-1). For all these reasons, it is not conclusive that the small tuffaceous lens taken as a reference for the dating of LOM3, and which does not occur on the LOM 3 section, is 3.31 ± 0.02 Ma. We stress that we are not claiming that this is not the age of the deposit, but simply that Harmand *et al.* did not convincingly report that such was the case.	Ignored.	None.	

Authoritative assertions have nothing to do with real science. The fact that five experienced geologists have determined that the assemblage must be *in situ* without further efforts to demonstrate it does not improve the poor scientific status of the currently held interpretations of the site. Science only knows about arguments and it is important to stress that such arguments must be empirically supported. However, these are still absent at Lomekwi 3. As cleverly stated by Fastovsky and Weishampel (2016, p. 217) when referring to unjustified exaggerated estimations of dinosaur body masses: “Bragging rights have little relationship to science”.

The truth is that in a period in which palaeoanthropology, like most other academic disciplines, is subjected to the tyranny of production, stasis is not acceptable. Researchers (especially young ones) and on-going projects need to show novelty, to show that knowledge progresses. Funding and job opportunities are tied to visibility (and unfortunately the growing pressure for mass media visibility of only exceptional results conditions the quality of many investigations). One must be innovative, come up with new ideas, new discoveries. The time was ripe in academia to accept a pre-*Homo* stone tool using and meat-eating small-brained hominin. Primate archaeologists had been shaping the mood for years. This is why when Dikika and Lomekwi 3 showed up, this was taken as confirmation of a whole set of preconceived ideas. If the discoveries were made by consolidated teams, which included major academic authorities (less likely to be questioned) and who belonged in prestigious academic institutions (which made them even less likely to be questioned), whose visibility in meetings and academic networks was prominent, all this would lead to an easy landing of the discoveries and the paradigm-changing implications that they had. Power has also something to do with it. We have been approached by senior scientists who also claimed scepticism about Lomekwi 3, but decided not to voice it for fear that it might jeopardize their research interests in East Africa. Politics come across science.

Lomekwi shows that academic dynamics have a life on their own independent of the scientific method. It also shows that paradigms can be created and maintained prior to solid scientific support. Even if eventually Lomekwi is vindicated by new more secure discoveries, that will not affect the description of this academic *modus operandi*. In the meantime, a whole paradigm was questioned and a new one emerged. This new interpretation now leaves the triggers and the biological basis of the new behaviour unexplained. More opportunities for research have just opened, if palaeoanthropologists can manage to keep their foundations on the ground.

In 2002, a series of probable footprints were found at the locality of Trachilos (Crete, Greece), that were argued to have belonged to a bipedal primate. The footprints were found on a 5.7 Ma (range = 8.5–5.6 Ma) sedimentary rock right before the Messinian salinity crisis (when the Mediterranean Sea had dried up). By the symmetry and size of the footprints it seems that they may have belonged to a bipedal creature. It was argued that the big toe seems to be aligned with the other toes, instead of occurring separate and in an oblique angle (like the opposable toes in chimpanzees and gorillas used for climbing). The footprints apparently showed a pentadactyl foot with laterally-decreasing toe size. One form also purportedly showed a strong ball impression from the hallux, that created a somewhat wider front foot aspect than modern human feet. The authors described the discovery as: “The print morphology suggests that the trackmaker was a basal member of the clade Hominini, but as Crete is some distance outside the known geographical range of pre-Pleistocene hominins we must also entertain the possibility that they represent a hitherto unknown late Miocene primate that convergently evolved human-like foot anatomy” (Gierliński *et al.*, 2017). The discovery is unexpected, and controversial because it is certainly ambiguous, but also potentially relevant for human evolution; hence the auction of the authors. They were aware “of the challenges of making such inferences when the implications run counter to conventional views on human evolution“. Two identifiable trackways were studied and the absence of prints attributable to the frontlimbs suggested a bipedal creature as the maker. The morphology differs from the human foot in the sole being “proportionately shorter, with a narrow tapering heel, and lacks a permanent arch”. Despite this, a geometric morphometric analysis of a set of landmarks clustered the Trachilos footprints with those of Laetoli and closer to modern humans than to any other primate. Regardless of the several issues that these ichnites raise as actual footprints, they do deserve consideration by the palaeoanthropological community for their potential and for assessing current paradigms on where and how human evolution was triggered.

There have been some claims that Miocene European hominids may have used bipedalism before African hominins. The recent discovery of the 11 Ma *Danuvius guggenmosi* is the latest of such claims (Böhme *et al.*, 2019). Previous claims of bipedality (and even hominin status) were made of *Rudapithecus* and *Graecopithecus* (Begun, 2015). Therefore, a bipedal primate in Miocenic Crete would not be too far-fetched. However, despite having much better information about context and age than Lomekwi 3, Trachilos remains virtually invisible (Crompton, 2017) in palaeoanthropological academia, and has certainly not made it (not even at the conjecture level) to the palaeoanthropology introductory books. The long-held paradigm that bipedalism emerged in Africa at the end of the Miocene is still too strong to be challenged; among other reasons, because bipedalism-derived features are still essential for identifying what a hominin is. If the door to the hominin realm is opened to European hominids, then certainty about what is a hominin can easily vanish. This is partly the reason why the Trachilos footprints were so difficult to reach publication in addition to their controversial nature as ichnites. The authors tried for years but “had a devil of a time getting this data published” (in words of one of the authors). We are not endorsing Trachilos as what the authors say they are, but they presented their arguments in a standard palaeoanthropological manner and need to be given the opportunity of critique on exactly the same grounds.

We will never know if the success of Trachilos had been different had the publishing authors been in a more consolidated palaeoanthropological team. Chances might also have been different had the paradigm being in the process of morphing (as was the case with Pliocene tool use). Examples of ideas and paradigms remaining staunchly immobile abound in palaeoanthropology. Let´s take the famous hunting-scavenging debate. Partisans of early hominin scavenging still survive because they are well settled academically. Their survival depends on continuously ignoring the large array of arguments and empirical evidence against the scavenging interpretation of early humans (see for example: Domínguez-Rodrigo, 2009, 2015).

DID POST-MODERNS GET THIS RIGHT?

One of the most influential post-modern thinkers, Michel Foucault, argued that there was a tight link between knowledge and power. He connected with hermeneutics by claiming that scientific knowledge was established around official discourses aiming at maintaining specific policies. He clearly established a link between power and academia, since knowledge emerges from the latter. Foucault (1975) believed that power was inherent to institutions more than to individuals. He created the figure of the “Panopticon”, as a model to explain how power emerges in a disindividualized system and automatizes its reproduction. Although he and other contemporaries argued that neoliberalism had sequestered academia, today this is more evident than when this critique was launched several decades ago. Today science has become more global, more corporative and more tightly linked through an intricate web of networks. This has expanded the power dynamics by creating a hierarchy where certain people have major decision power over funding resources, over what is accepted for publication and, more interestingly, over what ideas are promoted or demoted. This is done regardless of empirical support, in the margins of scientific heuristics. This creates trends that are promoted and followed by sheer interests of well-positioned individuals, through established networks. Such a process also results in trends that stagnate knowledge progress. In a book entitled "Fashion, Faith and Fantasy in the new physics of the universe", a prominent cosmologist and physicist, Roger Penrose (2017) shows that even in disciplines as purportedly objective as physics this process has been at work for a long time and that this could be detrimental to science:

"Indeed, ideally, it would be very reasonable to assert that such influences as fashion, or faith, or fantasy ought to be totally absent from the attitude of mind of those seriously dedicated to searching for the foundational underpinnings of our universe. Nature herself, after all, surely has no serious interest in the ephemeral whims of human fashion. Nor should science be thought of as a faith, the dogmas of science being under continual scrutiny and subject to the rigours of experimental examination, to be abandoned the moment that a convincing conflict arises with what we find to be the actuality of nature. And fantasy is surely the province of certain areas of fiction and entertainment, where it is not deemed essential that significant regard be paid to the requirements of consistency with observation, or to strict logic, or even to good common sense. Indeed, if a proposed scientific theory can be revealed as being too much influenced by the enslavement of fashion, by the unquestioning following of an experimentally unsupported faith, or by the romantic temptations of fantasy, then it is our duty to point out such influences, and to steer away any who might, perhaps unwittingly, be subject to influences of this kind […] A further point needs to be made here, with regard to research in theoretical physics that may be fashionable, yet far from what is plausible as a description of the world – indeed, as we shall find, often being in fairly blatant contradiction with current observations".

If several of the trendy ideas about modern physics (*i.e.*, string theory, multiverse theory) still need to be empirically supported, such ideas and theories can still be confidently subjected to demarcation criteria. Others, despite being at odds with empirical evidence are still maintained. Could it be because there is a network of scientists whose survival depends on it? Millionaire investments have been made on structures (such as colliders) that were built upon a set of theories. Such an evaluation has already been advanced by some (López-Corredoira, 2013; Unzicker & Jones, 2013).

The pervasive influence of power in academia has been identified through academic praxis:

"Without question, academia is a nexus of power. It is not only the power to form critique, but also the power to dictate certain forms of critique and knowledge, to establish norms of what can be said and how. Academic discourse is framed along these norms […] As such, the power of academia is the power to sort out, the power to allow to speak or to exclude from discussion. University is a place for powerful distinctions, and so the structures and power relations of academic discourse must be criticized" (Froebus, 2019, p. 38).

Constructive and deconstructive critique must, therefore, be subjected to logic and not be constrained by institutional origin. The underlying consequence of Froebus´ statement is that no serious critique will be created and spread within academia, because it would contradict its operational principles:

"Under the liberal conditions of a neoliberal university, academia is not only the place for critical thought and discourse, but it is also produced by and in academia –it is part of the neoliberal government(ality) of the university" (Froebus, 2019, p. 33).

The logical chain that ensues is: a) science is contingent in the sense that operates within a social construct and a historical moment; b) academic relations are constrained by power networks and a neoliberal system, more concerned about productivity than knowledge generation; c) there are academic elites, whose influence on the scientific praxis is paramount because of their control on resources, academic trends, and research publication; d) these elites are concentrated in economically-powerful nations, thereby showing the correlation between international economies & policies and academia; e) this latter situation shows the unavoidable link between political power and academic power (*i.e.*, between political elites and knowledge generation). In palaeoanthropology, the last points are clearly seen in the marginalization of non-Western academics; non-white African palaeoanthropologists, for example, are virtually invisible. This structure has also a perverse gender side. Most of the prominent palaeoanthropologists are also predominantly males.

This description of academic dynamics has long been known by several postmodern philosophers, despite the situation of self-denial exhibited by most scientists. Palaeoanthropology, being a non-experimentally based discipline, constructed around descriptive and comparative methods, with rather limited epistemological background, is a perfect example of this knowledge-biased generative process. Many of our interpretations are in essence not scientifically (*i.e.*, non-empirically) based. Trends emerge, evolve and disappear via morphing or replacement. Critical experiments do not exist in our discipline. Critical discoveries (although presented to the media as such) do not either. A common grasp of empirical support is description of sheer facts. The limited epistemological background of most palaeoanthropologists prevents them from articulating these facts efficiently in hypothesis-based theories. In the end, many of our interpretations are just well-intended tales, whose heuristics are unknown.

CONCLUSIONS

With palaeoanthropologists that overlook the epistemological heuristics of empirical evidence or use it partially by fitting it to their interpretive frameworks, interpretations become plural since they emerge from diverse conceptual universes. This renders all explanations relative. As in every other relative scenario, academics in this situation talk across each other. This may be seen by some as a sign of a healthy democratic academic dynamic. However, this apparent democracy has to be celebrated at a prize. If all interpretations are considered equally valid, because there is no need to abide by a scientific method, then one gives up any boundary between science, pseudo-science and other fantasy-derived approaches to purported knowledge. Without *demarcation* there is no science; which is the same as saying no certainty in knowledge. This democratic co-existence of interpretations also comes with another drawback. The interpretations made by research groups that hold a position of power will tend eventually to monopolize others. If interpretations persist not because of their heuristics, but because of the influence of those who hold them, academics will be reduced to politics. And as in politics, some countries will be more influential than others. As in politics, the impact of non-western countries will be minimized, the role of women will be more limited and the role of minorities will be reduced to quotas. This can actually be taken as a definition of the status of the representation of individuals in palaeoanthropology today (ironically, freedom would come with demarcation, with science: anybody should have the ability to defend or reject an interpretation based solely on its heuristics and lack of empirical contradictions. U.S. Ivy-League institutions should not be granted any departing advantage over interpretations produced by scientists in other less prestigious institutions, only because their social prestige adds a premium over the heuristics of knowledge production).

The non-experimental nature of palaeoanthropology makes its potential as a scientific discipline also rather limited. Hypothetico-deductive methods (needed for hypothesis-testing) require empirical testing. A potential solution to this could be the combination of hypothesis testing with probabilistic induction as Hempel suggested. Other alternatives can also be explored. The comparative method and the widespread historical concept of palaeoanthropology have hampered the perception that there are other ways of doing science. Unfortunately, both have enabled the asymmetries at all levels with which palaeoanthropology is practised today. Those alternatives that minimize the impact of politics, of power and client-based networks, as well as the pernicious effects of the neo-liberal hyper-productive dynamic that has permeated the discipline, leading to the questionable quality of "fast science", will ultimately do us a favor to get closer to the reality of our evolutionary history. Although the method is important, whether palaeoanthropology is a science or not ultimately depends on its practitioners.

The case discussed here for palaeoanthropology is not unique, but it is exemplary in that it makes objectivity harder than usual because it intertwines the object of research with the researcher (subject). Since ancient times in which humans were placed as the epitome of creation and even in more recent times in which taxonomy maintained human apex position in organic evolution (with a lack of modesty we qualified ourselves as *sapiens*), the current point in the evolutionary timeline calls for a more humble and less anthropocentric stance. However, we are still far from that, because in contrast, we have unashamedly proposed a new geological period bearing our name (the Anthropocene), whose duration is inferior in three orders of magnitude to the briefest geological time accepted by the IUGS. We understand that the progress in scientific knowledge –and the mere business that the publication of research has become through private publishers and global communication media- deserve an in-depth reflection on how the scientific method is being put into practice and abused today in academic praxis.

Supplementary information. No supplementary information.

Author contributions. Both authors have contributed equally to the work.

Competing interests. The authors declare that they do not have any competing interests.

Funding. The preparation of this paper did not require special funding.

Author details. **Manuel Domíngez-Rodrigo**[1,2,3], & **Luis Alcalá**[4]. [1]Institute of Evolution in Africa (IDEA), University of Alcalá de Henares, Madrid, Spain; [2]Area of Prehistory, Department of History and Philosophy, University of Alcalá de Henares, Alcalá de Henares, Spain; [3]Department of Anthropology, Rice University, 6100 Main St., Houston, TX 77005-1827, USA; m.dominguez.rodrigo@gmail.com; [4]Parque de las Ciencias, Andalucía-Granada, Av. de la Ciencia, s/n, 18006 Granada, Spain; alcala@parqueciencias.com

Acknowledgements. We thank Dr Carlos Martínez Pérez (Chief Editor SJP), Dr Ángela D. Buscalioni and an anonymous reviewer, as well as Manuel Alcaraz, for their insightful comments on an earlier version of this manuscript.

REFERENCES

Akerlof, G. A., & Michaillat, P. (2018). Persistence of false paradigms in low-power sciences. *Proceedings of the National Academy of Sciences of the United States of America, 115*(52), 13228–13233. doi: 10.1073/pnas.1816454115

Almécija, S., Tallman, M., Alba, D. M., Pina, M., Moyà-Solà, S., & Jungers, W. L. (2013). The femur of *Orrorin tugenensis* exhibits morphometric affinities with both Miocene apes and later hominins. *Nature Communications, 4*, 2888. doi: 10.1038/ncomms3888

Archer, W., Aldeias, V., & McPherron, S. P. (2020). What is 'in situ'? A reply to Harmand *et al.* (2015). *Journal of Human Evolution, 142*, 102740. doi: 10.1016/j.jhevol.2020.102740

Beebe, J. R., & Dellsén, F. (2020). Scientific Realism in the Wild: An Empirical Study of Seven Sciences and History and Philosophy of Science. *Philosophy of Science, 87*(2), 336–364. doi: 10.1086/707552

Begun, D. R. (2015). *The Real Planet of the Apes: A New Story of Human Origins*. Princeton University Press.

Beyries, S. (1988). *Industries lithiques: tracéologie et technologie*. British Archaeological Reports.

Binford, L. R. (1973). Interassemblage variability: The Mousterian and the "Functional" argument. In C. Renfrew (Ed.), *The Explanation of Culture Change: Models in Prehistory* (pp. 227–254). Duckworth.

Binford, L. R. (2002). *In Pursuit of the Past: Decoding the Archaeological Record*. University of California Press.

Binford, L. R. (2014). *Bones: Ancient Men and Modern Myths*. Academic Press.

Binford, L. R., & Binford, S. R. (1966). A Preliminary Analysis of Functional Variability in the Mousterian of Levallois Facies. *American Anthropologist, 68*(2), 238–295. doi: 10.1525/aa.1966.68.2.02a001030

Blumenschine, R. J. (1986). *Early hominid scavenging opportunities: implications of carcass availability in the Serengeti and Ngorongoro ecosystems*. British Archaeological Reports.

Böhme, M., Spassov, N., Fuss, J., Tröscher, A., Deane, A. S., Prieto, J., Kirscher, U., Lechner, T., & Begun, D. R. (2019). A new Miocene ape and locomotion in the ancestor of great apes and humans. *Nature*, *575*(7783), 489–493. doi: 10.1038/s41586-019-1731-0

Bordes, F. (1961). Mousterian Cultures in France. *Science, 134*(3482), 803–810. doi: 10.1126/science.134.3482.80

Brunet, M., Guy, F., Pilbeam, D., Mackaye, H. T., Likius, A., Ahounta, D., Beauvilain, A., Blondel, C., Bocherens, H., Boisserie, J.-R., De Bonis, L., Coppens, Y., Dejax, J., Denys, C., Duringer, P., Eisenmann, V., Fanone, G., Fronty, P., Geraads, D., Lehmann, T., Lihoreau, F., Louchart, A., Mahamat, A., Merceron, G., Mouchelin, G., Otero, O., Peláez Campomanes, P., Ponce de León, M., Rage, J.-C., Sapanet, M., Schuster, M., Sudre, J., Tassy, P., Valentin, X., Vignaud, P., Viriot, L., Zazzo, A., & Zollikofer, C. (2002). A new hominid from the Upper Miocene of Chad, Central Africa. *Nature, 418*(6894), 145–151. doi: 10.1038/nature00879

Bunge, M. (1998). *Philosophy of Science: From problem to theory*. Transaction Publishers.

Bunge, M. (1999). *Social Science Under Debate: A Philosophical Perspective*. University of Toronto Press.

Bunge, M. (2015). *Chasing reality: Strife over realism*. University of Toronto Press.

Bunge, M. (2017). *Philosophy of Science: Volume 1, From Problem to Theory*. Routledge.

Crompton, R. H. (2017). Making the case for possible hominin footprints from the Late Miocene (c. 5.7 Ma) of Crete? *Proceedings of the Geologists' Association. Geologists' Association, 128*(5–6), 692–693. doi: 10.1016/j.pgeola.2017.09.001

Curnoe, D. (2003). Problems with the use of cladistic analysis in palaeoanthropology. *Homo, 53*(3), 225–234. doi: 10.1078/0018-442X-00048

de la Torre, I., & Mora, R. (2014). The transition to the Acheulean in East Africa: an assessment of paradigms and evidence from Olduvai Gorge (Tanzania). *Journal of Archaeological Method and Theory, 21*(4), 781–823. doi: 10.1007/s10816-013-9176-5

De Renzi, M. (2005). Si la paleontología es una ciencia, debe ir más allá de las rutinas mecánicas y de las búsquedas de tesoros. In G. Meléndez, C. Martínez-Pérez, S. Ros, H. Botella, & P. Plasencia (Eds.), *Miscelánea Paleontológica* (pp. 3–45). SEPAZ, 6.

Diéguez, A. (1998). *Realismo científico. Una introducción al debate actual en la filosofía de la ciencia*. Universidad de Málaga.

Diéguez, A. (2020). *Transhumanismo. La búsqueda tecnológica del mejoramiento humano*. Herder Editorial.

Domínguez-Rodrigo, M. (1999). Flesh availability and bone modifications in carcasses consumed by lions: palaeoecological relevance in hominid foraging patterns. *Palaeogeography, Palaeoclimatology, Palaeoecology, 149*(1), 373–388. doi: 10.1016/S0031-0182(98)00213-2

Domínguez-Rodrigo, M. (2009). Are all Oldowan Sites Palimpsests? If so, what can they tell us about Hominid Carnivory? In E. Hovers, & D. R. Braun (Eds.), *Interdisciplinary Approaches to the Oldowan Vertebrate Paleobiology and Paleoanthropology* (pp. 129–147). Springer Netherlands.

Domínguez-Rodrigo, M. (2012). *Stone Tools and Fossil Bones: Debates in the Archaeology of Human Origins*. Cambridge University Press.

Domínguez-Rodrigo, M. (2015). Taphonomy in early African archaeological sites: Questioning some bone surface modification models for inferring fossil hominin and carnivore feeding interactions. *Journal of African Earth Sciences, 108,* 42–46. doi: 10.1016/j.jafrearsci.2015.04.011

Domínguez-Rodrigo, M., & Alcalá, L. (2016). 3.3-Million-year-old stone tools and butchery traces? More evidence needed. *PaleoAnthropology, 2016*, 46–53. doi: 10.4207/PA.2016.ART99

Domínguez-Rodrigo, M., & Alcalá, L. (2019). Pliocene Archaeology at Lomekwi 3? New Evidence Fuels More Skepticism. *Journal of African Archaeology, 17*(2), 173–176. doi: 10.1163/21915784-20190006

Domínguez-Rodrigo, M., Alcalá, L., & Luque, L. (2009). *Peninj: A research project on human origins 1995-2005*. Oxbow Books.

Domínguez-Rodrigo, M., Pickering, T. R., & Bunn, H. T. (2010). Configurational approach to identifying the earliest hominin butchers. *Proceedings of the National Academy of Sciences of the United States of America, 107*(49), 20929–20934. doi: 10.1073/pnas.101371110

Fang, F. C., & Casadevall, A. (2011). Retracted science and the retraction index. *Infection and Immunity, 79*(10), 3855–3859. doi: 10.1128/iai.05661-11

Fang, F. C., Steen, R. G., & Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences of the United States of America, 109*(42), 17028–17033. doi: 10.1073/pnas.1212247109

Fastovsky, D. A., & Weishampel, D. B. (2016, reprinted 2017). *Dinosaurs. A Concise Natural History*. Third ed. Cambridge University Press.

Feyerabend, P. (1993). *Against Method*. Verso.

Foucault, M. (1975). *Surveiller et punir: Naissance de la prison*. Gallimard.

Froebus, K. (2019). Critical perspectives on knowledge production in academia: introductory remarks. In A. N. Ahmad, M. Fielitz, J. Leinus, & M. Schlichte (Eds.), *Knowledge, normativity and power in academia: critical interventions. Normative Orders* (pp. 29–42). Campus Verlag.

Galik, K., Senut, B., Pickford M., Gommery, D., Treil, J., Kuperavage, A. J., & Eckhardt, R. B. (2004). External and internal morphology of the BAR 1002'00 *Orrorin tugenensis* femur. *Science, 305*(5689), 1450–1453. doi: 10.1126/science.1098807

Gallotti, R., & Mussi, M. (2018). The Emergence of the Acheulean in East Africa: Historical Perspectives and Current Issues. In R. Gallotti, & M. Mussi M (Eds.), *The Emergence of the Acheulean in East Africa and Beyond: Contributions in Honor of Jean Chavaillon* (pp. 1–12). Springer International Publishing.

Gibbons, A. (2004). Paleoanthropology. Oldest human femur wades into controversy. *Science*, *305*(5692), 1885.

Gidna, A. O., Kisui, B., Mabulla, A., Musiba, C., & Domínguez-Rodrigo, M. (2014). An ecological neo-taphonomic study of carcass consumption by lions in Tarangire National Park (Tanzania) and its relevance for human evolutionary biology. *Quaternary International, 322–323,* 167–180. doi: 10.1016/j.quaint.2013.08.059

Gierliński, G. D., Niedźwiedzki, G., Lockley, M. G., Athanassiou, A., Fassoulas, C., Dubicka, Z., Boczarowski, A., Bennett, M. R., & Ahlberg, P. E. (2017). Possible hominin footprints from the late Miocene (c. 5.7 Ma) of Crete? *Proceedings of the Geologists' Association, 128*(5–6), 697–710. doi: 10.1016/j.pgeola.2017.07.006

Gingerich, P. (1977). Correlation of tooth size and body size in living hominoid primates, with a note on relative brain size in *Aegyptopithecus* and *Proconsul*. *American Journal of Physical Anthropology, 47*(3), 395–398. doi: 10.1002/ajpa.1330470308

Gould, S. J. (1980). The promise of paleobiology as a nomothetic, evolutionary discipline. *Paleobiology, 6*(1), 96–118.

Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B 205*(1161), 581–598. doi: 10.1098/rspb.1979.0086

Grabowski, M., Hatala, K. G., & Jungers, W. L. (2018). Body mass estimates of the earliest possible hominins and implications for the last common ancestor. *Journal of Human Evolution, 122*, 84–92. doi: 10.1016/j.jhevol.2018.05.001

Grabowski, M., Hatala, K. G., Jungers, W. L., & Richmond, B. G. (2015). Body mass estimates of hominin fossils and the evolution of human body size. *Journal of Human Evolution, 85,* 75–93. doi: 10.1016/j.jhevol.2015.05.005

Hacking, I. (1982). Experimentation and Scientific Realism. *Philosophical Topics,13*(1), 71–87.

Haile-Selassie, Y. (2001). Late Miocene hominids from the Middle Awash, Ethiopia. *Nature, 412*(6843), 178–181.

Harcourt-Smith, W. E. H. (2007). The origins of bipedal locomotion. *Handbook of Paleoanthropology, 3,* 1483–1518. doi: 10.1007/978-3-642-39979-4_48

Harmand, S., Lewis, J. E., Feibel, C. S., Lepre, C. J., Prat, S., Lenoble, A., Boës, X., Quinn, R. L., Brenet, M., Arroyo, A., Taylor, N., Clément, S., Daver, G., Brugal, J.-P., Leakey, L., Mortlock, R. A., Wright, J. D., Lokorodi, S., Kirwa, C., Kent, D. V., & Roche, H. (2015). 3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya. *Nature, 521*(7552), 310–315. doi: 10.1038/nature14464

Harmand, S., Lewis, J. E., Taylor, N., Feibel, C. S., Boës, X., Prat, S., & Roche, H. (2019). Reply to Domínguez-Rodrigo and Alcalá: Interpretation Without Accurate Evidence Is Fantasy. *Journal of African Archaeology, 17*(2), 177–181. doi: 10.1163/21915784-20190007

Harvey, P. H., & Pagel. M. D. (1991). *The Comparative Method in Evolutionary Biology*. Oxford University Press.

Haslam, M., Hernández-Aguilar, A., Ling, V., Carvalho, S., de la Torre, I., DeStefano, A., Du, A., Hardy, B., Harris, J., Marchant, L., Matsuzawa, T., McGrew, W., Mercader, J., Mora, R., Petraglia, M., Roche, H., Visalberghi, E., & Warren, R. (2009). Primate archaeology. *Nature, 460*(7253), 339–344. doi: 10.1038/nature08188

Hempel, C. G. (1942). The Function of General Laws in History. *Journal of Philosophy, 39*(2), 35–48.

Hempel, C. G. (1966). *Philosophy of Natural Science*. Upper Saddle River, Prentice-Hall.

Hempel, C. G. (2001). *The Philosophy of Carl G. Hempel: Studies in Science, Explanation, and Rationality*. Oxford University Press.

Kuhn, T. S. (2012). *The Structure of Scientific Revolutions*. University of Chicago Press, 50th Anniversary Edition.

Lakatos, I. (1978). *The methodology of scientific research programmes. Philosophical Papers. Volume 1*. Cambridge University Press.

Leakey, M. D. (1971). *Olduvai Gorge: Volume 3, Excavations in Beds I and II, 1960-1963*. Cambridge University Press.

Lewis, J. E., & Harmand, S. (2016). An earlier origin for stone tool making: implications for cognitive evolution and the transition to *Homo*. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 371,* 20150233. doi: 10.1098/rstb.2015.0233

Lieberman, D. E., Pearson, O. M., Polk, J. D., Demes, B., & Crompton, A. W. (2003). Optimization of bone growth and remodeling in response to loading in tapered mammalian limbs. *Journal of Experimental Biology, 206*(18), 3125–3138. doi: 10.1242/jeb.00514

Lieberman, D. E., Polk, J. D., & Demes, B. (2004). Predicting long bone loading from cross-sectional geometry. *American Journal of Physical Anthropology, 123*(2), 156–171. doi: 10.1002/ajpa.10316

López-Corredoira, M. L. (2013). *The Twilight of the Scientific Age*. BrownWalker Press.

McPherron, S. P., Alemseged, Z., Marean, C. W., Wynn, J. G., Reed, D., Geraads, D., Bobe, R., & Béarat, H. A. (2010). Evidence for stone-tool-assisted consumption of animal tissues before 3.39 million years ago at Dikika, Ethiopia. *Nature, 466*(7308)*,* 857–860. doi: 10.1038/nature09248

Moyà-Solà, S., Köhler, M., & Rook, L. (1999). Evidence of hominid-like precision grip capability in the hand of the Miocene ape *Oreopithecus*. *Proceedings of the National Academy of Sciences of the United States of America, 96*(1)*,* 313–317. doi: 10.1073/pnas.96.1.313

Niiniluoto, I. (1999). *Critical Scientific Realism*. Oxford University Press.

Niiniluoto, I. (2012). *Truthlikeness*. Springer Science & Business Media.

Niiniluoto, I. (2013). *Is Science Progressive?* Springer Science & Business Media.

Nunn, C. L. (2011). *The Comparative Approach in Evolutionary Anthropology and Biology*. University of Chicago Press.

Pearson, O. M., & Lieberman, D. E. (2004). The aging of Wolff's 'law': ontogeny and responses to mechanical loading in cortical bone. *American Journal of Physical Anthropology, 125*(S39), 63–99. doi: 10.1002/ajpa.20155

Penrose, R. (2017). *Fashion, Faith, and Fantasy in the New Physics of the Universe*. Princeton University Press.

Percival, C. J., & Richtsmeier, J. T. (2017). *Building Bones: Bone Formation and Development in Anthropology*. Cambridge University Press.

Pickford, M. (2018). Piping, a geomorphological process relevant to African palaeontology and archaeology: sedimentary, taphonomic and biostratigraphic implications. *Communications to the Geological Survey of Namibia, 20,* 59–86.

Pobiner, B. L. (2015). New actualistic data on the ecology and energetics of hominin scavenging opportunities. *Journal of Human Evolution, 80*, 1–16. doi: 10.1016/j.jhevol.2014.06.020

Pobiner, B. L. (2020). The zooarchaeology and paleoecology of early hominin scavenging. *Evolutionary Anthropology, 29*(2)*,* 68–82. doi: 10.1002/evan.21824

Pontzer, H., Lieberman, D. E., Momin, E., Devlin, M. J., Polk, J. D., Hallgrímsson, B., & Cooper, D. M. L. (2006). Trabecular bone in the bird knee responds with high sensitivity to changes in load orientation. *The Journal of Experimental Biology, 209*(1)*,* 57–65. doi: 10.1242/jeb.01971

Popper, K. (2002). *The logic of scientific discovery*. Routledge Classics New York.

Proffitt, T., Luncz, L. V., Falótico, T., Ottoni, E. B., de la Torre, I., & Haslam, M. (2016). Wild monkeys flake stone tools. *Nature, 539*(7627)*,* 85–88. doi: 10.1038/nature20112

Ramírez, M. J. (2007). *Homology as a parsimony problem: a dynamic homology approach for morphological data*. *Cladistics, 23,* 588–612. doi: 10.1111/j.1096-0031.2007.00162.x

Rodríguez, L., Jiménez-Arenas, J. M., & De Renzi, M. (2010). Estudio de la variabilidad craneal en *Homo sapiens* e implicaciones para la taxonomía de los homininos fósiles: variación geográfica y clinas. In A. Santos, M. Mayoral, G. Meléndez, C. Marques da Silva, & M. Cachão (Eds.), *III Congreso Ibérico de Paleontología. XXVI Jornadas de la Sociedad Española de Paleontología* (pp. 253–256). SEPAZ, 9.

Rook, L., Bondioli, L., Köhler, M., Moyà-Solà, S., & Macchiarelli, R. (1999). *Oreopithecus* was a bipedal ape after all: evidence from the iliac cancellous architecture. *Proceedings of the National Academy of Sciences of the United States of America, 96*(15)*,* 8795–8799. doi: 10.1073/pnas.96.15.8795

Rudwick, M. J. S. (1972). *The Meaning of Fossils. Episodes in the History of Palaeontology*. Macdonald, London and American Elsevier.

Ruff, C. B., Holt, B., & Trinkaus, E. (2006). Who's afraid of the Big Bad Wolff?: ''Wolff's Law'' and Bone Functional Adaptation. *American Journal of Physical Anthropology, 129*(4)*,* 484–498. doi: 10.1002/ajpa.20371

Russo, G. A., & Shapiro L. J. (2013). Reevaluation of the lumbosacral region of *Oreopithecus bambolii*. *Journal of Human Evolution, 65*(3)*,* 253–265. doi: 10.1016/j.jhevol.2013.05.004

Sahle, Y., El Zaatari, S., & White, T. D. (2017). Hominid butchers and biting crocodiles in the African Plio–Pleistocene. *Proceedings of the National Academy of Sciences of the United States of America, 114*(50)*,* 13164–13169. doi: 10.1073/pnas.1716317114

Sánchez-Yustos, P., Diez-Martín F., Díaz, I., Fraile, C., Uribelarrea, D., Mabulla, A., Baquedano, E., & Domínguez-Rodrigo, M. (2019). What comes after the Developed Oldowan B debate? Techno-economic data from SHK main site (Middle Bed II, Olduvai Gorge, Tanzania). *Quaternary International, 526,* 67–76. doi: 10.1016/J.QUAINT.2019.07.034

Semaw, S., Renne, P., Harris, J. W., Feibel, C. S., Bernor, R. L., Fesseha, N., & Mowbray, K. (1997). 2.5-million-year-old stone tools from Gona, Ethiopia. *Nature, 385*(6614)*,* 333–336. doi: 10.1038/385333a0

Semaw, S., Rogers, M. J., Simpson, S. W., Levin, N. E., Quade, J., Dunbar, N., McIntosh, W. C., Cáceres, I., Stinchcomb, G. E., Holloway, R. L., Brown, F. H., Butler, R. F., Stout, D., & Everett, M. (2020). Co-occurrence of Acheulian and Oldowan artifacts with *Homo erectus* cranial fossils from Gona, Afar, Ethiopia. *Science Advances, 6*(10)*,* eaaw4694. doi: 10.1126/sciadv.aaw4694

Smith, R. J., & Wood, B. (2017). The principles and practice of human evolution research: Are we asking questions that can be answered? *Comptes Rendus Palevol, 16*(5)*,* 670–679. doi: 10.1016/j.crpv.2016.11.005

Smith, R. J., Albrecht, G. H., Damuth, J., Di Bacco, M., Fortelius, M., Gingerich, P. D., Godfrey, L. R., Sutherland, M. R., Jungers, W. L., Leigh, S. R., Leney, M. D., Foley, R. A., Leonard, W. R., Robertson, M. L., Leutenegger, W., McHenry, H. M., Martin, R. D., Pilbeam, D., Plavcan, J. M., Wheeler, P. E., Wood, B. A., & Collard, M. (1996). Biology and Body Size in Human Evolution: Statistical Inference Misapplied [and Comments and Reply]. *Current Anthropology, 37*(3)*,* 451–481.

Smolin, L. (2007) *The Trouble with Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next*. Houghton Mifflin Company.

Smolin, L. (2013). *Time Reborn: From the Crisis in Physics to the Future of the Universe*. Houghton Mifflin Harcourt.

Tixier, J. (1996). Technologie et typologie: dérives et sclérose. *Quaternaria Nova, 6,* 15–21.

Unzicker, A., & Jones, S. (2013). *Bankrupting Physics: How Today's Top Scientists are Gambling Away Their Credibility*. St. Martin´s Press.

Uribelarrea, D., Martín-Perea, D., Díez-Martín, F., Sánchez-Yustos, P., Domínguez-Rodrigo, M., Baquedano, E., & Mabulla, A. (2017). A reconstruction of the paleolandscape during the earliest Acheulian of FLK

West: The co-existence of Oldowan and Acheulian industries during lowermost Bed II (Olduvai Gorge, Tanzania). *Palaeogeography, Palaeoclimatology, Palaeoecology, 488,* 50–58. doi: 10.1016/j.palaeo.2017.04.014

Wallace, I. J., Burgess, M. L., & Patel, B. A. (2020). Phalangeal curvature in a chimpanzee raised like a human: Implications for inferring arboreality in fossil hominins. *Proceedings of the National Academy of Sciences of the United States of America, 117*(21), 11223–11225. doi: 10.1073/pnas.2004371117

Wallace, I. J., Demes, B., Mongle, C., Pearson, O. M., Polk, J. D., & Lieberman, D. E. (2014). Exercise-Induced Bone Formation Is Poorly Linked to Local Strain Magnitude in the Sheep Tibia. *PLoS ONE, 9*(6), e99108. doi: 10.1371/journal.pone.0099108

White, T. D., Asfaw, B., Beyene, Y., Haile-Selassie, Y., Lovejoy, C. O., Suwa, G., & WoldeGabriel, G. (2009). *Ardipithecus ramidus* and the paleobiology of early hominids. *Science, 326*(5949), 75–86. doi: 10.1126/science.1175802

Wolpoff, M. H., Hawks, J., Senut, B., Pickford, M., & Ahern, J. (2006). An ape or the age: is the Toumaï TM 266 cranium a hominid? *PaleoAnthropology, 2006*, 36–50.

Wolpoff, M. H., Senut, B., Pickford, M., & Hawks, J. (2002). Palaeoanthropology. *Sahelanthropus* or '*Sahelpithecus*'? *Nature, 419*(6907), 581–582. doi: 10.1038/419581a

Wood, B., & Harrison, T. (2011). The evolutionary context of the first hominins. *Nature, 470*(7334), 347–352. doi: 10.1038/nature09709