

**The value of epistemic disagreement in scientific practice.
The case of *Homo floresiensis***

This is the version of a paper as it was accepted after peer review in *Studies in History and Philosophy of Science, Part A*.

<http://www.sciencedirect.com/science/article/pii/S0039368113000162>

<http://dx.doi.org/10.1016/j.shpsa.2013.02.002>

Please cite this paper as follows: De Cruz, H. & De Smedt, J. (2013). The value of epistemic disagreement in scientific practice. The case of *Homo floresiensis*. *Studies in History and Philosophy of Science Part A*, 44, 169–177.

Helen De Cruz
Institute of Philosophy
University of Leuven
K Mercierplein 2, 3000 Leuven, Belgium
helen.decruz@hiw.kuleuven

Faculty of Philosophy, University of Oxford
Woodstock Road OX2 6GG Oxford, United Kingdom

Johan De Smedt
Department of philosophy and ethics, Ghent University
Blandijnberg 2, 9000 Gent, Belgium
johan.desmedt@ugent.be

Abstract

Epistemic peer disagreement raises interesting questions, both in epistemology and in philosophy of science. When is it reasonable to defer to the opinion of others, and when should we hold fast to our original beliefs? What can we learn from the fact that an epistemic peer disagrees with us? A question that has received relatively little attention in these debates is the value of epistemic peer disagreement—can it help us to further epistemic goals, and, if so, how? We investigate this through a recent case in paleoanthropology: the debate on the taxonomic status of *Homo floresiensis* remains unresolved, with some authors arguing the fossils represent a novel hominin species, and others claiming that they are *Homo sapiens* with congenital growth disorders. Our examination of this case in the recent history of science provides insights into the value of peer disagreement, indicating that it is especially valuable if one does not straightaway defer to a peer's conclusions, but nevertheless remains open to a peer's evidence and arguments.

Keywords: Peer disagreement, *Homo floresiensis*, paleoanthropology, historical sciences, prehistory.

1. Disagreement in scientific practice

What should we do when faced with people who disagree with us, despite having access to the same evidence, and being similar in cognitive capacities? This question captivates a sizable portion of the current social epistemology literature. Some authors

(e.g., Christensen, 2007; Feldman, 2007) argue in favor of a *conciliatory position*, according to which it is unreasonable to hold fast to one's beliefs in the face of peer disagreement. Others (e.g., van Inwagen, 1999; Rosen, 2001) maintain a *steadfast position*: it reasonable to hold fast to one's prior belief in the face of dissenting epistemic peers. There are also mixed positions, for instance, Lackey (2010) recommends revising one's belief in the face of disagreement, except if it enjoys a very high degree of justified confidence. The burgeoning literature on epistemic peer disagreement has relied primarily on toy examples, such as bill calculation (people in a restaurant add up the bill and come to different sums), horse race (two equally well-sighted, well-positioned spectators disagree on which horse crossed the finish line first), and elementary math (two adults argue about whether the outcome of $2 + 2 = 4$ or 5).

Discussions on the significance of disagreement in philosophy of science predate those in social epistemology by several decades. Philosophers of science have examined whether, and under what circumstances, it is better on balance to defer to a rival scientific theory, and when it is recommendable to stick to one's guns. In contrast to the toy examples in epistemology, philosophers of science (e.g., Kitcher, 1995) typically consider historical cases, such as the shift from geocentrism to heliocentrism in early modern astronomy, or the rivalry between oxygen and phlogiston theories of combustion. They also formulate analytic and computer models to explore the value of maintaining or resolving disagreements (e.g., De Cruz & De Smedt, 2012; Kitcher 1990; Weisberg & Muldoon, 2009). This growing empirically-informed literature indicates that there is no simple rule of thumb for handling disagreement within a scientific field. Indeed, often a combination of several strategies is optimal at the level of the scientific community. For example, Weisberg and Muldoon (2009) found that mixed populations, composed of mavericks who stubbornly pursue unpopular research strategies and conformists who follow the most successful strategies, outperform pure populations composed of either mavericks or conformists. More generally put: a heterogeneous community of scientists hedges its bets, and can achieve a fruitful division of labor, whereas a community where everyone follows the most successful and promising hypotheses fares less well in the long run (Kitcher, 1995, chapter 8). In this respect, philosophy of science moved away from the classical Kuhnian view of normal science as a monolithic entity, emphasizing instead the productive role of "the active interplay of tenaciously held views" in everyday scientific practice (Feyerabend, 1970, 209).

This focus on the level of the scientific community tells us relatively little about what individual reasoners should do in the face of disagreement. For even if it benefits the scientific community at large, maintaining disagreement may be suboptimal for individual practitioners. For example, Priestley may have been unreasonable to hold on to the phlogiston theory as long as he did, even though his work turned out to be valuable for chemistry as a discipline (Kitcher, 1990, 7). Kitcher (1995, 344) juxtaposes these levels: he argues that a community of clear-headed, rational scientists, does less well than a community of sullied, stubborn scientists. Also, it remains unclear what scientists should do if two hypotheses H and H^* both seem equally promising. Some philosophers of science (e.g., Turner, 2007, 47) have argued that the rational response in this case is to suspend judgment until more evidence becomes available, or simply to move on to more tractable research questions. We propose that maintaining disagreement is valuable also for individual scientists, because it helps them to achieve epistemic goals (e.g., to acquire more evidence). If this is indeed the case, it provides a rational motivation to stick to one's beliefs in the face of dissenting peers.

The aim of this paper is to combine views on disagreement from philosophy of science and social epistemology to study the value of peer disagreement in scientific practice. We will focus on a contemporary case study from the historical sciences: the debate on the taxonomic status of *Homo floresiensis*, a small hominin from the Indonesian island of Flores. We will first examine whether, and if so, how the notion of epistemic peer from social epistemology can be translated into scientific practice. We then consider epistemological issues that arise in the historical sciences, and how this can give rise to disagreement between epistemic peers. We will specify three ways in which maintaining epistemic peer disagreement is valuable for furthering epistemic goals, and argue that these benefits are available to a far lesser extent if one defers to an epistemic peer's opinion or suspends judgment.

2. Who are epistemic peers in scientific practice?

For our discussion, it is important to get a clear notion of what is meant by epistemic peers and by disagreement in a scientific setting. At first sight, this seems like a straightforward task, as the term 'peer' is frequently used in scientific practice, for instance, in open peer commentary and peer review. However, in such cases, the parties in question are not always truly epistemic peers: a full professor with a 25-year track record and over 100 papers in peer-reviewed journals who referees a manuscript by a beginning graduate student is engaged in peer review, but she clearly is the student's epistemic superior. On the other hand, the idealized concept of epistemic peer as adopted in the epistemological literature seems overtly restrictive to be applicable to science. First, scientists are often confronted with multiple, mutually incompatible alternative explanations and opinions regarding the same phenomenon, and not just with two positions¹. Second, scientists rarely arrive at their position in isolation from each other, whereas idealized cases in social epistemology typically involve judgments that have been formed entirely independently. Third, whereas toy examples presuppose an unambiguous and comprehensive body of evidence and principles that are shared between peers, this cannot be assumed in scientific practice, where there can be disagreement about what counts as relevant evidence and which principles can be used to arrive at it. If equal evidence were required for peer disagreement, virtually no scientific case would fall under this label, with perhaps the exception of members of the same lab, studying under the same supervisor, working on a specific and narrowly defined problem. However, restricting epistemic peers to clean cases makes the social epistemological toolbox of marginal relevance for scientific practice (and indeed, for any real-world situation) where messy and complex cases predominate. Therefore, we will seek to adopt a broader notion of epistemic peer than is customary in social epistemology, a notion that captures actual scientific practice, and that nevertheless remains pertinent for social epistemologists.

The term epistemic peer, as originally conceived by Gutting (1982), described epistemic peers in terms of intellectual virtues such as attentiveness, intelligence and thoroughness. In recent discussions, the meaning of this term has shifted: two agents A and B are epistemic peers with regard to a given question Q if they are *cognitive equals* (Lackey, 2010), i.e., possess similar cognitive capacities and are subject to similar cognitive limitations, and if they are *evidential equals* (Christensen, 2007), i.e., have access to the same evidence with respect to Q.

¹ A telling example is the debate on the reasons for the fall of the Western Roman empire among archaeologists, historians and classicists—a debate where over 500 theories are on offer (Randsborg, 1989, 230).

Cognitive equality means that all parties are knowledgeable in disciplines relevant to Q, and that they have similar levels of expertise. Although it will be in practice impossible to find scientists with an identical track record, a roughly similar quality of scientific output (as assessed by, e.g., citation metrics, the quality of journals published in) seems a reasonable approximation. The most interesting cases of scientific peer disagreement occur when the disagreeing parties are all experts in a domain D that is relevant for investigating Q. Experts are people who possess extensive pools of knowledge in D, as well as a set of skills to deploy this knowledge successfully to new questions within D (Goldman, 2001, 92).

Evidential equality is a more complex issue. The highly idealized requirement of epistemic peers who have access to exactly the same evidence (as put forward by e.g., Feldman, 2007) is rarely, if ever met, in real-life situations. In a narrow sense, the evidence is what is immediately relevant to Q, and what is equally accessible to all parties, such as published research findings (e.g., descriptions of fossils in refereed papers). In a broad sense, the evidence consists of someone's total body of background knowledge, including but not limited to that obtained through training and earlier research. This latter type of evidence is not easily shared. For example, primatologists may disagree about the nature of social cognition in primates. Researchers from the Kyoto Primate Research Institute (e.g., Matsuzawa, 2009) believe that chimpanzee cultural transmission is relevantly similar to human cultural transmission, whereas those from the Max Planck Institute for Evolutionary Anthropology (e.g., Tomasello & Herrmann, 2010) think that it is very different. This disagreement arises partly through differences in evidence in the broad sense, such as procedural know-how on setting up experiments, approaching primates during these experiments, and how to house and treat them. Kyoto primates are housed in family groups, often composed of mothers with infants, and maintain close proximity with human investigators, whereas Max Planck primates are housed in a zoo context, interact more with individuals of the same age, and have less close relationships with their human experimenters. The Kyoto primatologists think proximity to primates is the best way to obtain results, whereas Max Planck researchers believe this decreases objectivity. According to de Waal (2003) these divergent approaches find their roots in the distinct religious legacies in Japan and Germany, with Buddhism stressing continuity between humans and other animals (humans can reincarnate as animals and vice versa), whereas western culture, with its Christian legacy, sees humans and animals as radically distinct. If these worldviews exert enduring influence on methodology, evidence in the broad sense becomes expansive indeed.

In the light of this, it may be better to restrict evidential equality to the narrow sense. Evidence in the narrow sense for scientific practice consists mainly of the hard published data and other primary source material that is relevant to Q. While evidence in the broad sense can influence one's interpretation of evidence in the narrow sense, this latter type of evidence is discussed and shared in a way that one expects in epistemic peers. Although the interpretation of evidence in the narrow sense may diverge, e.g., Japanese researchers appeal to concepts like 'master apprenticeship' (e.g., Matsuzawa, 2009) and other human forms of cultural learning, whereas Western researchers may stress individual learning processes (e.g., Tomasello & Herrmann, 2010), the experimental results are there for all to see. Explaining the findings of rival groups will typically appeal to shared theoretical views on chimpanzee psychology. Thus, having access to evidence in the narrow sense seems like a reasonable approximation for evidential equality. By these criteria, many scientists will be each

other's epistemic peers due to the high degree of professionalization that characterizes scientific fields.

Disagreement means that scientists hold contrary or incompatible beliefs. The simplest case is when one believes that p and another that $\neg p$, but the term can be broadened to any case in which scholars hold incompatible doxastic attitudes toward the same proposition. Scientific peer disagreement is typically subject to what Lynch (2010) has termed *deep epistemic disagreement*. In such cases, parties have the same epistemic goals, but they disagree on which epistemic principles should be used to achieve them. Epistemic goals can be distinguished from pragmatic goals, in that they are not primarily directed at bringing about changes in the world (e.g., seeking food and shelter), but at bringing about changes in our cognitive attitude toward the world (e.g., getting to know the location of shops in your neighborhood). In the context of scientific research, epistemic goals are directed at bringing about true beliefs in a particular domain of enquiry D and to avoid false beliefs in D . Epistemic principles are normative principles that say that some sources or ways of forming a belief are reliable, i.e., likely to produce true beliefs. Examples include methods like radiocarbon dating, DNA amplification, or computer simulations, and modes of inference like inference to the best explanation, or consilience of multiple lines of evidence. Deep epistemic disagreement occurs when there are common epistemic goals, but there is no agreement on epistemic principles, and where there is no further principle or evidence, accepted by both parties, available to settle the disagreement.

3. A case study from paleoanthropology

To examine the value of peer disagreement, we focus on a contemporary case study in the historical sciences: the (still unresolved) discussion on the taxonomic status of *Homo floresiensis* in paleoanthropology. Paleoanthropology is a historical natural science. Historical sciences attempt to infer unobservables from the past (e.g., dinosaurs) by looking at present-day surviving clues (e.g., fossils). Paleoanthropology is the multidisciplinary study of extinct and extant hominins, combining principles and methods from among others paleontology, archaeology, primatology, and physical anthropology. Hominin evolution is examined using fossil hominin bones and other material evidence, including tools, archaeological traces of living sites, fossilized pollen, and animal bones. Over the past few decades, the extraction of DNA from fossil remains, combined with DNA analysis of extant humans and great apes has also brought genetic evidence into this picture. Given that humans have a deep-seated desire to know more about their origins, it is perhaps not surprising that paleoanthropology is a prestigious discipline: new hominin fossils meet with extensive media coverage, and frequently grace the covers of *Nature* and *Science*.

Paleoanthropology is subject to epistemological constraints of the historical sciences. The hominins palaeoanthropologists study are no longer directly observable. Competing hypotheses cannot be tested by predicting events in well-controlled laboratory conditions, but rather, are evaluated by their capacity to unify and explain a diversity of current observations. Consilience of multiple lines of evidence is an important guideline to decide between competing hypotheses (Cleland 2011, Turner, 2007).

Peculiar to paleoanthropology, compared to some other historical sciences like geology, is that theory-formation is heavily dependent on another discipline, mainstream evolutionary biology, especially when it comes to assigning fossil hominins to species and to proposing evolutionary lineages that link these species. For the most part, it has been descriptive rather than explicitly theoretical:

paleoanthropologists attempt to provide taxonomic classifications for the hominin fossil record, rather than develop novel theories on how these fossils connect in an evolutionary framework (Tattersall, 2000). Theory-formation is data-driven to the extent that a single convincing counterexample suffices to refute cherished theories. For example, the long-established view that *Homo ergaster* evolved anagenetically from *Homo habilis* has been successfully challenged by the find of fossil remains of *Homo habilis* and *Homo ergaster* in the same archaeological layer (Spoor et al., 2007).

This characterization of paleoanthropology is important to understand the context in which disagreement typically occurs within the field. Disagreements tend to arise when the interpretation of a new fossil is inconsistent with prevailing notions of the course of human evolution, as when Dart (1925) considered the so-called Taung child (*Australopithecus africanus*) to be a human ancestor rather than an ape. As we will see in the case of *Homo floresiensis*, the disagreement that underlies such cases can be deep, i.e., the disagreeing parties have the same epistemic goals, but diverge in their epistemic principles.

Our case study concerns the status of bones of a small hominin found in the cave of Liang Bua on the Indonesian island of Flores in 2003 (Brown et al., 2004). Due to the climatic conditions of the cave (humid and warm) the bones are not fossilized and have the consistency of wet blotting paper². Up to now, incomplete skeletal remains of 14 individuals have been found, the oldest of which are dated to about 95,000 BP³, the most recent to 18–16,000 BP. The majority of *Homo floresiensis* fossils were found in the youngest layer. *Homo sapiens* remains are found in younger levels of the cave, beginning at about 11,000 BP, separated from the older bones by a layer of volcanic ash (Morwood et al., 2005).

The type specimen LB1 is an almost complete skeleton of an adult of undetermined sex, dated to 18,000 BP. He or she stood only about 1.06 m tall, falling well outside of the variation of current human and extinct hominin populations. Even the tiny *Australopithecus afarensis* Lucy (3.8 million years ago) was somewhat taller than these hominins. At 400 cc, the brain volume of the Flores individuals is smaller than most known hominin fossils, only slightly larger than that of a chimpanzee (Martin et al., 2006). In October 2004, the newly discovered fossils received extensive media attention; the hobbits, as they became known, captured the interest of paleoanthropologists as well as the lay audience. Some paleoanthropologists disagreed with the discoverers' interpretation of the fossils. As a result, two mutually incompatible positions arose:

The insular dwarfism hypothesis (H_{id}) the remains are from a late, hitherto unknown, distinct hominin species, *Homo floresiensis*, which evolved locally on Flores from an early hominin ancestor due to insular dwarfism, in a manner similar to many other relatively large mammals that migrated to islands. Within H_{id} there are several views about the evolutionary trajectory of *Homo floresiensis* (it descended from *Homo erectus*, *Homo habilis*, or perhaps even a late australopithecine). All proponents of H_{id}

² Although the term 'subfossil' is technically more correct to describe ancient bones that have not (completely) fossilized, we will regularly use the term 'fossil' since it is also used for ancient, non-fossilized bones.

³ BP stands for 'Before Present'. In archaeology and paleoanthropology, it is the standard way to specify dates in the past. To avoid the problem of an ever-shifting present, 'Present' by convention refers to 1950.

agree that the speciation event leading to *Homo floresiensis* occurred early, between 1 and 2 million years ago.

The pathological modern human hypothesis (H_{mph}) the fossils are members of our own species, *Homo sapiens*, who suffered from a congenital disorder that involved a severe reduction in brain size (microcephaly) and body size (congenital dwarfism). Within H_{mph} there are several hypotheses about the disorder that led to these phenotypic traits. All proponents of H_{mph} agree that *Homo floresiensis* is a *Homo sapiens*, and hence the emergence of the *Homo floresiensis* phenotype was late and transient.

To determine the taxonomic status of fossils, paleoanthropologists analyze morphological features of skeletons and classify these as either *primitive* or *modern*. A primitive feature is one that resembles that of present-day apes. Examples include large teeth, prominent brow-ridges, long forearms and a receding forehead. Modern features are those that diverge from apes and resemble those of present-day humans, such as a large brain volume, small teeth and longer lower limbs. Like most hominins, *Homo floresiensis* exhibits a mix of primitive and modern traits, but there is disagreement about how this should be interpreted.

The original team of excavators (Brown et al., 2004) proposed that the fossils be classified as a new hominin species. Initially, they placed their find in a new genus, *Sundanthropus*, but the reviewers of *Nature* thought that the cranium had clear affinities with our own genus, so the name was changed to *Homo floresiensis* (Aiello, 2010). Dismissing microcephaly and congenital dwarfism, Brown et al. (2004) argued that LB1 exhibits a mosaic of primitive and modern features, and that it probably evolved locally from *Homo erectus*, a species that was present in Indonesia as early as 1.9 million years ago. Since *Homo erectus* was a large-brained hominin (1000–1200 cc), this interpretation implies that a small-brained species had descended from a large-brained ancestor, a pattern that was not previously observed in human evolution, where the general trend is increasing brain size over time (Rightmire, 2004). This anomaly, together with other features of the find, such as the recent dating and the presence of what were originally described as advanced stone tools, led some researchers to propose that the remains were not members of a distinct new hominin species, but rather pathological individuals of our own (e.g., Martin et al., 2006).

The fact that the excavators (Brown et al., 2004) of these remains were the first to propose H_{id} may be reason for suspicion. The discovery of a hitherto unknown hominin species leads to publications in prestigious journals and guarantees continued funding for future excavations. The taxonomy of a given hominin fossil, i.e., whether it belongs to a new species or not, seems to trump all other features of a find, including its preservation and quality⁴. There are both prestige-based and pecuniary interests that favor assigning new fossils to new species. All things being equal, this fact should lower our confidence in H_{id} (see Goldman, 2001, 104-105, for a discussion of the role of biases in disagreement between experts). Nonetheless, H_{id} is also endorsed by

⁴ This is dramatically illustrated by the media coverage that surrounded the discovery of a single finger bone of a hominin from southern Siberia dated to 30,000 BP (Krause et al., 2010) which had DNA that was distinct from both *Homo sapiens* and *Homo neanderthalensis*. Up to then, these two species were the only ones believed to inhabit Siberia, so the find represents an unknown hominin species.

paleoanthropologists who did not partake in the original excavations, including Lahr and Foley (2004), Gordon et al. (2008) and Aiello (2010).

H_{id} proponents identify primitive features in the morphology of *Homo floresiensis* to support their hypothesis. Tocheri et al. (2007) compared the wrist-bones of the Liang Bua remains with those of *Homo sapiens* and extant nonhuman apes. They argue that *Homo floresiensis* has primitive wrist-bones, supporting the view that these specimens evolved from a very early, not yet fully bipedal hominin ancestor, like *Homo habilis*. Falk et al. (2007) made scans of LB1's skull to examine its internal brain organization. They contend that it is different from that of present-day microcephalic individuals. Larson et al. (2007) examined the anatomical proportions of the well-preserved shoulder bones, and concluded that *Homo floresiensis* did not have a modern human shoulder configuration: the clavicle was relatively short, and the scapula was more protracted than that of modern humans. The shoulders are similar to those of early hominins, like the Nariokotome boy, an African *Homo ergaster* fossil dated to 1.5 million years ago. The authors stipulated that no known diseases or growth disruptions in modern humans could cause the shoulder morphology of *Homo floresiensis*, which supports H_{id}. The pelvis (Jungers et al., 2009c) looks unlike that of modern humans (including people affected with congenital dwarfism) and resembles that of australopithecines and other not fully bipedal early hominins. Jungers et al. (2009b) argue that the foot morphology is unlike that of modern humans, but rather exhibits a mosaic of primitive and modern features, in particular in the long, curved toes that differ markedly from the short, straight toes of modern humans⁵. To these authors, the primitive morphology of the foot-bones in the fossils indicates that *Homo floresiensis* evolved from an early hominin, perhaps even a late australopithecine. Taken together, this evidence gathered in support of H_{id} points to an early hominin as the most likely ancestor of *Homo floresiensis*. It evolved locally into its diminutive form due to selective pressures for smaller body and brain size that are typical for island populations. Such forms of insular dwarfism are not uncommon in mammals: Flores had dwarf stegodons, relatives of the elephants (Brumm et al., 2010), and fossils of dwarf elephants and dwarf hippopotami were found on Crete (Diamond, 1992).

Proponents of H_{pmh} base their assignment of the Liang Bua fossils to our species primarily on their small brain size. At 400 cc, the brain volume of the fossils is smaller than of most other known hominins (e.g., 600 cc for *Homo habilis*, one of the earliest members of our genus). Martin et al. (2006) calculate that the degree of brain size reduction is greater than what one would expect from insular dwarfism. The majority of dwarfed mammals have relatively large brain sizes because brains do not scale down to the same extent as the rest of the body. Congenital disorders that cause a severe reduction in brain volume occur in modern humans. Such forms of microcephaly sometimes leave cognitive functions intact, which would account for the presence of stone tools. Weber et al. (2005) argue that microcephalic individuals have a highly variable internal brain organization: the peculiar brain morphology of LB1 is thus not an argument against H_{pmh}. Microcephaly is not a natural kind, but rather an umbrella term for a wide range of growth disorders that result in brain reduction; it can be caused by many different syndromes. The scarcity of endocasts (casts of the inside of a skull) from microcephalics also makes it difficult to decide whether the internal

⁵ The shape of the modern human foot is the result of selective pressures associated with bipedalism. Hominins that have curved foot-bones, such as the late australopithecines, were probably less exclusively bipedal, combining bipedalism with other forms of locomotion, such as tree climbing.

brain structure of the LB1 skull could be the product of microcephaly. Because it is not a well-delineated syndrome, proponents of H_{pmh} need to specify which growth disorder produced the particular morphology of the fossils in order to have a testable claim. Hershkovitz et al. (2007) propose that the reduction in growth is due to Laron syndrome, a rare genetic disorder that blocks the production of growth hormones. People with Laron syndrome are similar in stature and brain size as *Homo floresiensis*. According to Obendorf et al. (2008), the humans from Flores had myxoedematous endemic cretinism, a local congenital condition that is not uncommon among historical and present-day Indonesian island populations. It affects the proper functioning of the thyroid gland, leading to severe dwarfism and microcephaly. They also argue that this condition gives rise to some of the features that appear primitive, such as the wrist-bone.

4. A case of deep epistemic disagreement

The parties in the discussion about the status of *Homo floresiensis* are epistemic peers. They have similar levels of training and experience that are relevant to paleoanthropology, all have published in prominent journals, all have access to evidence in the narrow sense, i.e., the published research findings and the fossils themselves. But their evidence in the broad sense differs. Indriati (cited in Aiello 2010, 171) has observed that most adherents to H_{id} come from a background in paleontology and archaeology, whereas the majority of proponents of H_{pmh} have a physical anthropological (clinical or geneticist) background⁶. The former concentrate on the morphology of the fossils and formulate their hypotheses on the basis of their analysis of primitive and modern features. The latter examine the fossils in the context of clinical conditions, and formulate clinical hypotheses to explain the *Homo floresiensis* morphology. This difference in evidence in the broad sense partly explains the disagreement between the two factions, despite evidential equality in the narrow sense.

At present, no evidence can settle the debate in a way that convinces all parties. Cleland (2002, 480, 482) terms evidence that adjudicates between several competing hypotheses a “smoking gun”: a smoking gun is “a trace (or subcollection of traces) that (so-to-speak) cinches the case for a particular causal story.” For *Homo floresiensis*, there are several possible smoking guns. If DNA had been extracted, and if it turned out to match the *Homo sapiens* genome, this would be a decisive victory for H_{pmh} . If, on the other hand, it turned out to lie outside of the normal human variability, the paleoanthropological community would probably come to endorse H_{id} . Unfortunately, in contrast to many other recent fossils, such as those of *Homo neanderthalensis*, the bones contain no extractable DNA, as the humid and warm conditions on Flores have been detrimental to DNA preservation⁷. A decisive form of proof in favor of H_{id} would be the find of transitional fossils that significantly predate the arrival of *Homo sapiens* in Asia, documenting the evolution to the *Homo floresiensis* morphology. Unfortunately, although stone tools of 1 million years old were found on Flores (Brumm et al., 2010), hominin fossils were lacking in these layers. By contrast, a pathological condition in anatomically modern humans that presents all of the features

⁶ However, this division is not absolute, for example Teuku Jacob, who had a paleontological background, adhered to H_{pmh} , and Dean Falk, with a clinical background in neuroanatomy, is a proponent of H_{id} .

⁷ Typically, ancient DNA can only be recovered if the remains are preserved in a cold and dry environment. The Max Planck Institute, which previously successfully extracted Neanderthal DNA failed to extract DNA from the Liang Bua remains.

observed in *Homo floresiensis*, and not just a subset of them, would decisively tilt the balance in favor of H_{pmh} . Currently, however, no pathology can explain the entire *Homo floresiensis* phenotype.

As is characteristic for deep disagreement, the differing epistemic principles shape the way both parties perceive and evaluate the same body of evidence. Evidence in the narrow sense is on the table for all epistemic peers, both in the form of detailed descriptions in peer-reviewed papers and in access to the actual fossils and tools. The way this evidence is interpreted, however, depends on the epistemic principles the paleoanthropologists adhere to. It turns out that almost every morsel of evidence concerning *Homo floresiensis* can be interpreted in two mutually conflicting ways. For example, the face of LB1 exhibits considerable bilateral asymmetry. Brown et al. (2004) argue that this asymmetry is a post-mortem result of the pressure of the archaeological layers on the remains—given that the bones were soft and unfossilized, it is not unlikely that the skull got distorted. For Jacob et al. (2006), however, this asymmetry is evidence that the skull belonged to a diseased individual.

The dating between 95,000 to 16,000 BP also seems open to at least two interpretations. Those who argue for H_{id} contend that the older specimens (between 95,000 and 75,000 BP) significantly predate the oldest *Homo sapiens* fossils in Indonesia (Culotta, 2007), and that the youngest specimens (16,000 BP) are still significantly older than the oldest attested *Homo sapiens* on Flores at 11,000 BP (Morwood et al., 2005). By contrast, adherents to H_{pmh} maintain that the dating is consistent with their interpretation, since anatomically modern *Homo sapiens* evolved 195,000 BP in Africa, and left that continent by 120,000 BP. Indeed, *Homo sapiens* fossils dating to at least 100,000 BP were found in southern China (Liu et al., 2010).

The observation that Liang Bua cave contains only remains of very small individuals is seen in a different light through each hypothesis. Seen through the lens of H_{pmh} the fact that all the skeletons are small is explained as the result of a religious practice that led to the burial of all microcephalic individuals in a single site (Martin et al., 2006). On the other hand, none of the skeletons show signs of intentional burial, grave gifts or ochre (Culotta, 2007), which are typical for prehistoric anatomically modern human burials. This supports H_{id} since lack of formal burial is almost universal in premodern hominins, except perhaps in Neanderthals.

The interpretation of the tools found in association with *Homo floresiensis* also exhibits significant theory-ladenness. Given that many forms of microcephaly (especially those proposed in H_{pmh}) only have modest effects on cognitive function, one would expect an advanced toolkit. Indeed, H_{pmh} proponents like Jacob et al. (2006) and Martin et al. (2006) describe the stone tools as modern and sophisticated, noting the presence of blades⁸. By contrast, proponents of H_{id} (e.g., Moore, 2007) note that most of the *Homo floresiensis* stone tools are simple, produced by direct percussion, and that the blades can be explained as an incidental byproduct of their knapping technique.

5. What does disagreement in the *Homo floresiensis* case accomplish?

We will now turn to the question of whether or not proponents of H_{id} and H_{pmh} should maintain their disagreement, or whether suspension of judgment is a more adequate response, by focusing on the value of disagreement. Turner (2005, 2007, 47) has argued that in cases where two rival hypotheses H and H^* are empirically equivalent,

⁸ Blades are elongated, relatively small stone tools, associated with *Homo sapiens* in Africa since about 70,000 BP.

scientists should suspend judgment. However, what we typically see in actual scientific practice (as in this case) is that scientists—even those not originally involved with the analysis of the fossils, like Aiello (2010)—tend to pick sides. But this is a factual observation; it tells us little about whether it is a good thing to maintain disagreement in the face of a dissenting peer with differing epistemic principles. If continued disagreement is valuable in helping to further epistemic goals (in this case, acquiring true beliefs about the taxonomic status of the Flores remains), this seems *prima facie* a good reason to hold fast to one's position. In what follows, we suggest three interrelated advantages of epistemic peer disagreement: the generation of new evidence, the reassessment of existing evidence, and overcoming confirmation bias. We will illustrate each of these with elements from the *Homo floresiensis* debate.

5.1 The generation of new evidence

The extensive discussions and commentaries in peer-reviewed journals indicate that the disagreeing paleoanthropologists recognize each other as epistemic peers. They treat disagreement itself as evidence that needs to be considered, and in response to it attempt to find new evidence to strengthen their original case. As a consequence, scientific evidence increases as a direct result of the disagreement. This is not merely by providing an extra motivation to work harder, but, crucially by giving an incentive to uncover evidence that not only confirms one's original hypothesis, but that also manages to convince dissenting peers that one's original hypothesis is true. When successful, this newly-generated evidence may provide a smoking gun. A smoking gun not only confirms an existing hypothesis, but may also disconfirm rival hypotheses (Jeffares, 2008). Without competing hypotheses, historical scientists may (erroneously) believe they have settled the case, abandoning the search for further evidence. The presence of several competing hypotheses helps to guide scientists to look for additional evidence that will be relevant to all parties.

Historical scientists are more constrained than experimental scientists in uncovering new evidence. They are subject to the historical contingencies of what has been preserved and what has been destroyed over time. This led Turner (2005, 2007) to conclude that the historical sciences are epistemically inferior to the experimental sciences, where one can generate evidence in controlled experimental conditions. But crucially, historical events leave many traces, *most* of which remain unexamined. For example, it is standard archaeological practice to only dig up a tiny portion of an archaeological site, saving time and resources, and leaving material for future archaeologists (with perhaps better techniques) to investigate. Many traces are initially overlooked, even though they constitute relevant evidence, for example, microscopic patterns of wear on stone tools reveal what the tools were used for (e.g., slicing meat or cutting wood). Other traces may at present not be seen as evidence, but can become evidence once appropriate techniques become available to study them. For example, methods to extract ancient DNA from fossils, developed in the later decades of the 20th century, made DNA sequencing of extinct organisms possible. While there is an undeniable stochastic element in historical scientific practice, disagreement can foster a search for evidence that not only confirms one's hypothesis in the light of one's background beliefs, but that also provides a disconfirmation of the other party's hypothesis in the light of their background views.

In the case of *Homo floresiensis* the primitive shape of the wrist-bone is an observation that, according to H_{id} proponents, confirms the hypothesis that these fossils are members of a distinct hominin species (e.g., Brown et al., 2004). Proponents of H_{pmh} (Obendorf et al., 2008) counter this claim with the observation that the primitive

wrist-bone can be found in some congenital growth disorders, such as myxoedematous endemic cretinism. Proponents of H_{id} (Jungers et al., 2009a) respond to this with a comparison between the *Homo floresiensis* fossils and individuals affected by cretinism, arguing that the cranial capacity of LB1 falls outside of the variation observed in adults with myxoedematous endemic cretinism, and that the ratio of the foot to femur length is far higher in *Homo floresiensis* individuals than in contemporary humans with this disorder.

The interpretation of the evidence is theory-dependent. Since it seems plausible that paleoanthropologists are aware that the disagreement is deep, i.e., they and their dissenters also disagree about the epistemic principles through which evidence is gathered, they will plausibly also realize that the other party will not be overtly impressed by their newly gathered evidence. As a result, there has been an effort to find evidence that is not dependent on the hypotheses in question. Take the presence of one-million-year-old stone tools on Flores (Brumm et al., 2010). This observation is theory-independent with respect to H_{id} and H_{pmh} because it does not rely on hypotheses concerning the taxonomic status of the hominin remains. Rather, it is dependent on other theories in paleoanthropology and its ancillary disciplines, for example, theories concerning radioactive decay on which geochronological dating methods are based. Nevertheless, these ancient tools could be taken as evidence that confirms H_{id} because, according to this hypothesis, *Homo floresiensis* is descended from an earlier hominin species after a long period of isolation, and this hypothesis can only work if Flores was indeed occupied by hominins during a considerable time. Theories on which dating methods are based are firmly established within paleoanthropology (and also in other sciences, including geology and chemistry, that form the basis of geochronological dating methods). They are also more widely endorsed by paleoanthropologists than the hypotheses about *Homo floresiensis*, which are tentative and concern a highly localized phenomenon. For one thing, proponents of H_{id} do not agree among themselves on which ancestor *Homo floresiensis* evolved from, *Homo erectus*, *Homo habilis*, or a late australopithecine⁹. Paleoanthropologists are more deeply committed to theories underlying their dating methods—part and parcel of their discipline—than they are to theories on species classification of particular hominin fossils.

Another example of an observation that is not dependent upon the theories that are being tested is the recent find of small-bodied human remains on the island of Palau, Micronesia (Berger et al., 2008). These *Homo sapiens* are dated to 2,900–900 BP. The individuals were small in body size (0.94 m–1.20 m), even when compared to other pygmoid populations. Their brain size (about 1000 cc) falls at the very low end of the range in humans, but is still more than twice the size of that of *Homo floresiensis*. More importantly, however, like *Homo floresiensis*, the Palau skeletons possess primitive features. This indicates that insular dwarfism can occur in modern humans, and raises the possibility that the hominins from Flores, rather than pathological individuals, are normal *Homo sapiens* subject to extreme insular dwarfism. If this explanation could account for the *Homo floresiensis* morphology, one could postulate a new hypothesis that combines elements from H_{id} and H_{pmh} —the individuals from Liang Bua would indeed be subject to insular dwarfism, but they would not be ancient hominins, but present-day *Homo sapiens*. Thus, evidence that is

⁹ *Sundanthropus*, the name originally proposed for the species, seems a better designation than *Homo floresiensis*, since their opinion seems to shift toward late australopithecine. After all, australopithecines are not members of the genus *Homo*, and cladistic considerations dictate that species that do not descend from early *Homo* cannot be assigned to the genus *Homo*.

not theory-laden with respect to the competing hypotheses can generate novel ideas that are not colored by the rival hypotheses. This type of evidence is more likely to be generated when there is epistemic conflict than when there is assent, because if the theories under scrutiny are never called into question, evidence that is independent from existing hypotheses tends to get overlooked. Sustaining a disagreement leads to an increase in both the quantity of evidence and its quality.

5.2 Re-evaluation of existing evidence and assumptions

Peer disagreement not only leads to uncovering new evidence, but may also direct scientists to a closer scrutiny of existing evidence and assumptions, which would otherwise not have been re-examined. As a result of the *Homo floresiensis* debate, researchers have paid closer attention to the question of how the mammalian brain scales during severe body reduction. Initially, the relationship between brain and body mass was seen as a simple and unproblematic linear scaling relationship where *Homo floresiensis* does not fit into, which was taken as evidence against H_{id} . For example, Martin (2006) argued that microcephaly was the best explanation for the morphology and proportions of the *Homo floresiensis* cranium, because it did not scale as expected. However, more recent work kindled by this debate (e.g., Weston & Lister, 2009) shows that the actual relationship is more complex in animals like dwarfed hipopotami, which suggests that the scaling of the *Homo floresiensis* brain might not be problematic after all. Vanucci et al. (2011) took a closer look at the scaling of brains of microcephalic individuals, and compared these to LB1, *Homo sapiens* and *Homo erectus*. They found that the scaling of LB1's cranium falls outside of the range of present-day microcephalics.

According to Cleland (2011), in the historical sciences causes are locally overdetermined by phenomena in the present. For example, hominin occupation leaves several material traces, including stone tools, butchery marks on animal bones, traces of hearths, and hominin fossils. The presence of only a subset of these archeological features is sufficient to determine that hominins were present. On the other hand, traces of historical processes tend to degrade over time (e.g., stone tools survive longer than bones in most contexts), and in some cases the remaining traces make it impossible to decide between competing theories. Turner (2007) therefore argues that even if the future overdetermines the past, this has no epistemic consequences, because of the pervasive destruction of evidence. Cleland and Turner disagree on whether traces epistemically underdetermine or overdetermine their causes. Cleland (2002) believes that historical events leave so many traces that these overdetermine an event. Turner (2007) holds that the deterioration of traces is so pervasive that we are left with several competing theories that explain the same set of traces, leaving us unable to decide between them. We think that both authors overstate their position, in part due to their choice of case studies. Cleland focuses on clear-cut cases where a consensus is reached and the evidence overwhelmingly favors one hypothesis, as in an asteroid collision as the cause of the K-T mass extinction 65 million years ago. Turner on the other hand considers unresolved cases where disagreement reigns, e.g., the color of dinosaurs (which in the meantime has been resolved for several dinosaur species). Turner is correct in arguing that most work in the historical sciences does not provide clear-cut evidence in favor of one hypothesis as it does for the K-T mass extinction. However, he claims that there are plenty of examples in the historical sciences where we will never be able to adjudicate between competing hypotheses, *no matter what future evidence* will become available to us; they are “equally well supported by all the empirical evidence that will ever be available to us” (Turner, 2005, 217). This is a very

strong claim, especially in the light of the ingenuity of historical scientists to come up with new techniques and experimental procedures that allow them to analyze and sharpen attenuated traces. For example, careful analysis of taphonomic processes can allow researchers to recognize biases in the preservation of archaeological material.

The choice of these technologies and procedures is to an important extent informed by prior working hypotheses. Without the competing view that *Homo floresiensis* was a microcephalic *Homo sapiens*, there would have been little incentive to study the scaling of microcephalic brains, or indeed, to probe the scaling relationship between brain and body size in mammalian evolution. As this form of evidence in the narrow sense is relevant for both parties, disagreement has played a role in uncovering new techniques and models that aim to throw new light on existing evidence. As Niven (2006, 57) puts it: “The arguments raised by critics and advocates alike prompt a re-examination of ideas about what is possible during the evolution of the [mammalian] brain.” Historical scientists are not only interested in regularities (like the scaling of brains) in an instrumental sense. They are also interested in these regularities for their own sake; they are “actively interested in regularities across event types” (Jeffares, 2008, 474). Regularities in brain size evolution have been an enduring interest in paleoanthropology, as is evident by attempts to chart broad evolutionary patterns in hominin brain size evolution (e.g., Rightmire, 2004). As a result of the disagreement in the *Homo floresiensis* debate, an in-depth look at evidence of hominin brain evolution has yielded important new insights in the scaling between brain and body size in mammalian evolution.

5.3 An antidote to confirmation bias

We have argued that the debate on *Homo floresiensis* is subject to deep epistemic disagreement. Those who accept H_{id} find it prima facie likely that a novel hominin species could have evolved on Flores as a result of insular dwarfism. By contrast, proponents of H_{pmh} think it prima facie unlikely that such a species exists: “evolution [of a separate hominin species] over millennia in total isolation is unproved, unlikely” (Jacob et al., 2006, 13426). In their research, both parties tend to concentrate on evidence that is in agreement with their own hypothesis. A large body of evidence in the psychology of reasoning (see Mercier, 2010, for an overview) suggests that this attitude is pervasive: people exhibit confirmation bias, a tendency to selectively attend and evaluate evidence more positively that is in line with their prior beliefs. They also exhibit disconfirmation bias, i.e., they are more likely to dismiss evidence that disconfirms their own beliefs, or that confirms beliefs that are incompatible with their own convictions. Unlike some other biases, such as susceptibility to the conjunction fallacy, confirmation bias does not attenuate as a result of schooling; highly educated people are as susceptible to it as those with less academic training.

In line with this, psychological studies indicate that scientists are subject to confirmation bias when they are assessing and evaluating empirical evidence (e.g., Mahoney & DeMonbreun, 1977; Nickerson, 1998). In particular, scientists are prone to dismiss evidence that disconfirms their hypothesis on methodological grounds: Fugelsang et al. (2004) found that molecular biologists are reluctant to consider anomalous data as relevant evidence, dismissing them instead as errors or experimental artifacts.

It is revealing that the authors in the *Homo floresiensis* debate recognize this bias in their opponents, but fail to spot it in themselves: proponents of H_{pmh} were accused by their opponents of “cherry picking features and ignoring counterevidence,” whereas H_{id} advocates “largely looked for otherness” according to their skeptics (cited

in Culotta, 2006, 1028). Epistemic peer disagreement has the advantage that it forces one to pay attention to anomalous data that one initially failed to detect or had glossed over as a result of confirmation bias. Moreover, disconfirmation bias leads scientists to be critical and especially vigilant to their opponents' arguments and evidence. More generally, disagreement plays a crucial role in reasoning, as humans tend to reason better when they have to argue their case against opponents than when they reason in isolation (Mercier & Sperber, 2011). As we saw in the *Homo floresiensis* case, H_{id} proponents initially did not consider factors other than taxonomy that could result in a primitive wrist morphology (until clinical factors were pointed out to them), whereas defenders of H_{pmh} initially did not consider the morphological variation in microcephalic endocasts (until this was highlighted by their opponents), but relied mainly on brain size to assign *Homo floresiensis* to modern humans.

6. Conclusion

Epistemic peer disagreement is valuable because it brings about an increase in relevant evidence, a re-evaluation of existing evidence and assumptions, and a decrease in confirmation bias. These are instrumental in furthering epistemic goals, and can increase the probability that the taxonomic status of *Homo floresiensis* will be resolved. Maintaining a disagreement among epistemic peers is valuable for individual scientists: especially in dialogue with dissenting voices scientists will carefully check their hypotheses, look for new techniques and procedures to uncover new evidence that would also convince the other party, and try to make their case as convincing as possible.

References

- Aiello, L.C. (2010). Five years of *Homo floresiensis*. *American Journal of Physical Anthropology*, 142, 167–179.
- Berger, L.R., Churchill, S.E., De Klerk, B., & Quinn, R.L. (2008). Small-bodied humans from Palau, Micronesia. *PLoS One*, 3, e1780.
- Brown, P., Sutikna, T., Morwood, M. J., Soejono, R.P., Jatmiko, Saptomo, E.W., & Due, R.A. (2004). A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia. *Nature*, 431, 1055–1061.
- Brumm, A., Jensen, G.M., van den Bergh, G.D., Morwood, M.J., Kurniawan, I., Aziz, F., & Storey, M. (2010). Hominins on Flores, Indonesia, by one million years ago. *Nature*, 464, 748–752.
- Christensen, D. (2007). Epistemology of disagreement: The good news. *Philosophical Review*, 116, 187–217.
- Cleland, C.E. (2002). Methodological and epistemic differences between historical science and experimental science. *Philosophy of Science*, 69, 447–451.
- Cleland, C.E. (2011). Prediction and explanation in historical natural science. *British Journal for the Philosophy of Science*, 62, 551–582.
- Culotta, E. (2006). Skeptics seek to slay the 'hobbit,' calling Flores skeleton a modern human. *Science*, 313, 1028–1029.

- Culotta, E. (2007). The fellowship of the hobbit. *Science*, 317, 740.
- Dart, R.A. (1925). *Australopithecus africanus*: The man-ape of South Africa. *Nature*, 115, 195–199.
- De Cruz, H., & De Smedt, J. (2012). Evolved cognitive biases and the epistemic status of scientific beliefs. *Philosophical Studies*, 157, 411–429.
- de Waal, F.M. (2003). Silent invasion: Imanishi's primatology and cultural bias in science. *Animal Cognition*, 6, 293–299.
- Diamond, J.M. (1992). Twilight of the pygmy hippos. *Nature*, 359, 15.
- Falk, D., Hildebolt, C., Smith, K., Morwood, M. J., Sutikna, T., Jatmiko, Saptomo, E. Imhof, H., Seidler, H., & Prior, F. (2007). Brain shape in human microcephalics and *Homo floresiensis*. *Proceedings of the National Academy of Sciences USA*, 104, 2513–2518.
- Feldman, R. (2007). Reasonable religious disagreements. In L. Anthony (Ed.), *Philosophers without gods* (pp. 194–214). Oxford: Oxford University Press.
- Feyerabend, P. (1970). Consolations for the specialist. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 197–230). Cambridge: Cambridge University Press.
- Fugelsang, J.A., Stein, C.B., Green, A.E., & Dunbar, K.N. (2004). Theory and data interactions of the scientific mind: Evidence from the molecular and the cognitive laboratory. *Canadian Journal of Experimental Psychology*, 58, 86–95.
- Goldman, A.I. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological Research*, 63, 85–110.
- Gordon, A., Nevell, L., & Wood, B. (2008). The *Homo floresiensis* cranium (LB1): Size, scaling, and early *Homo* affinities. *Proceedings of the National Academy of Sciences USA*, 105, 4650–4655.
- Gutting, G. (1982). *Religious belief and religious skepticism*. Notre Dame, IN: University of Notre Dame Press.
- Hershkovitz, I., Kornreich, L., & Laron, Z. (2007). Comparative skeletal features between *Homo floresiensis* and patients with primary growth hormone insensitivity (Laron Syndrome). *American Journal of Physical Anthropology*, 134, 198–208.
- Jacob, T., Indriati, E., Soejono, R. P., Hsü, K., Frayer, D.W., Eckhardt, R.B., Kuperavage, A.J., Thorne, A., & Henneberg, M. (2006). Pygmoid Australomelanesian *Homo sapiens* skeletal remains from Liang Bua, Flores: Population affinities and pathological abnormalities. *Proceedings of the National Academy of Sciences USA*, 103, 13421–13426.

- Jeffares, B. (2008). Testing times: regularities in the historical sciences. *Studies in history and philosophy of science C*, 39, 469-475.
- Jungers, W.L., Falk, D., Hildebolt, C., Smith, K., Prior, F., Tocheri, M.W., Orr, C.M., Burnett, S.E., Larson, S.G., Djubiantono, T., & Morwood, M.J. (2009c). The hobbits (*Homo floresiensis*) were not cretins. *American Journal of Physical Anthropology Supplement*, 48, 161.
- Jungers, W.L., Harcourt-Smith, W.E.H., Wunderlich, R.E., Tocheri, M.W., Larson, S.G., Sutikna, T., Due, R.A., & Morwood, M.J. (2009b). The foot of *Homo floresiensis*. *Nature*, 459, 81–84.
- Jungers, W. L., Larson, S.G., Harcourt-Smith, W., Morwood, M.J., Sutikna, T., Due, R.A., & Djubiantono, T. (2009a). Descriptions of the lower limb skeleton of *Homo floresiensis*. *Journal of Human Evolution*, 57, 538–554.
- Kitcher, P. (1990). The division of cognitive labor. *Journal of Philosophy*, 87, 5–22.
- Kitcher, P. (1995). *The advancement of science. Science without legend, objectivity without illusions*. Oxford: Oxford University Press.
- Krause, J., Fu, Q., Good, J.M., Viola, B., Shunkov, M., Derevianko, A., & Pääbo, S. (2010). The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. *Nature*, 464, 894–897.
- Lackey, J. (2010). What should we do when we disagree? In T.S. Gendler & J. Hawthorne (Eds.), *Oxford Studies in Epistemology* (Vol. 3, pp. 274–293). Oxford: Oxford University Press.
- Lahr, M. & Foley, R. (2004). Human evolution writ small. *Nature*, 431, 1043–1044.
- Larson, S.G., Jungers, W.L., Morwood, M. J., Sutikna, T., Jatmiko, Saptomo, E., Due, R.A., & Djubiantono, T. (2007). *Homo floresiensis* and the evolution of the hominin shoulder. *Journal of Human Evolution*, 53, 718–731.
- Liu, W., Jin, C., Zhang, Y., Cai, Y., Xing, S., Wu, X., Cheng, H., Edwards, R. L., Pan, W., Qin, D., An, Z., Trinkaus, E., & Wu, X. (2010). Human remains from Zhirendong, South China, and modern human emergence in East Asia. *Proceedings of the National Academy of Sciences USA*, 107, 19201–19206.
- Lynch, M.P. (2010). Epistemic circularity and epistemic incommensurability. In A. Haddock, A. Millar, & D. Pritchard (Eds.), *Social epistemology* (pp. 262–277). Oxford: Oxford University Press.
- Mahoney, M.J. & DeMonbreun, B.G. (1977). Psychology of the scientist: An analysis of problem-solving bias. *Cognitive Therapy and Research*, 1, 229–238.

Martin, R.D., MacLarnon, A.M., Phillips, J.L., Dussubieux, L., Williams, P.R., & Dobyns, W.B. (2006). Comment on “The Brain of LB1, *Homo floresiensis*”. *Science*, 312, 999b.

Matsuzawa, T. (2009). The chimpanzee mind: In search of the evolutionary roots of the human mind. *Animal Cognition* 12: S1-S9.

Mercier, H. (2010). The social origins of folk epistemology. *Review of Philosophy and Psychology*, 1, 499–514.

Mercier, H. & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34, 57-74.

Moore, M.W. (2007). Lithic design space modelling and cognition in *Homo floresiensis*. In A.C. Schalley & D. Khlentzos (Eds.), *Mental states: Evolution, function, nature* (Vol. 1, pp. 11–33). Amsterdam: John Benjamins.

Morwood, M.J., Brown, P., Jatmiko, Sutikna, T., Saptomo, E.W., Westaway, K.E., Due, R.A., Roberts, R., Maeda, T., Wasisto, S., & Djubiantono, T. (2005). Further evidence for small-bodied hominins from the Late Pleistocene of Flores, Indonesia. *Nature*, 437, 1012–1017.

Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2, 175-220.

Niven, J.E. (2006). Brains, islands and evolution: Breaking all the rules. *Trends in Ecology and Evolution*, 22, 57–59.

Obendorf, P.J., Oxnard, C.E., & Kefford, B.J. (2008). Are the small human-like fossils found on Flores human endemic cretins? *Proceedings of the Royal Society B*, 275, 1287–1296.

Randsborg, K. (1989). The demise of Antiquity: Europe and the Mediterranean in the first millennium AD. *Annual Review of Anthropology*, 18, 227–244.

Rightmire, G.P. (2004). Brain size and encephalization in Early to Mid-Pleistocene *Homo*. *American Journal of Physical Anthropology*, 124, 109–123.

Rosen, G. (2001). Nominalism, naturalism, epistemic relativism. *Philosophical Perspectives*, 15, 69–91.

Spoor, F., Leakey, M.G., Gathogo, P.N., Brown, F.H., Antón, S.C., McDougall, I., Kiarie, C., Manthi, F.K., & Leakey, L.N. (2007). Implications of new early *Homo* fossils from Ileret, east of Lake Turkana, Kenya. *Nature*, 448, 688–691.

Tattersall, I. (2000). Paleoanthropology: The last half-century. *Evolutionary Anthropology*, 9, 2–16.

Tocheri, M.W., Orr, C.M., Larson, S.G., Sutikna, T., Jatmiko, Saptomo, E.H., Due, R.A., Djubiantono, T., Morwood, M.J., & Jungers, W. (2007). The primitive wrist of

Homo floresiensis and its implications for hominin evolution. *Science*, 317, 1743–1745.

Tomasello, M. & Herrmann, E. (2010). Ape and human cognition: What's the difference? *Current Directions in Psychological Science*, 19, 3–8.

Turner, D. (2005). Local underdetermination in historical science. *Philosophy of Science*, 72, 209–230.

Turner, D. (2007). *Making prehistory. Historical science and the scientific realism debate*. Cambridge: Cambridge University Press.

van Inwagen, P. (1999). Is it wrong everywhere, always, and for anyone to believe anything on insufficient evidence? In E. Stump & M.J. Murray (Eds.), *Philosophy of religion: The big questions* (pp. 273–284). Malden: Blackwell.

Vannucci, R.C., Barron, T.F., & Holloway, R.L. (2011). Craniometric ratios of microcephaly and LB1, *Homo floresiensis*, using MRI and endocasts. *Proceedings of the National Academy of Sciences USA*, 108, 14043–14048.

Weber, J., Czarnetzki, A., & Pusch, C. (2005). Comment on “The brain of LB1, *Homo floresiensis*”. *Science*, 310, 236b.

Weisberg, M., & Muldoon, R. (2009). Epistemic landscapes and the division of cognitive labor. *Philosophy of Science*, 76, 225–252.

Weston, E.M., & Lister, A.M. (2009). Insular dwarfism in hippos and a model for brain size reduction in *Homo floresiensis*. *Nature*, 459, 85–88.