[Forthcoming in *Studies in History and Philosophy of Science*, Part A. Accepted 17 August, 2015]

A Second Look at the Colors of the Dinosaurs

Derek D. Turner
Department of Philosophy and
Goodwin-Niering Center for the Environment
Connecticut College
270 Mohegan Avenue
New London, CT 06320, USA
Derek.turner@conncoll.edu

Abstract

In earlier work, I predicted that we would probably not be able to determine the colors of the dinosaurs. I lost this epistemic bet against science in dramatic fashion when scientists discovered that it is possible to draw inferences about dinosaur coloration based on the microstructure of fossil feathers (Vinther et al. 2008). This paper is an exercise in philosophical error analysis. I examine this episode with two questions in mind. First, does this case lend any support to epistemic optimism about historical science? Second, under what conditions is it rational to make predictions about what questions scientists will or will not be able answer?

Keywords: fossils, historical science, paleontology, underdetermination

1. Introduction. In earlier work, I predicted that we would probably not be able to determine the colors of the dinosaurs (Turner 2005, 2007). I lost this epistemic bet against science in dramatic fashion when scientists discovered that it is possible to draw inferences about dinosaur coloration based on the microstructure of fossil feathers (Vinther et al. 2008). A number of critics soon pointed out the failed prediction (Jeffares 2010, Stanford 2010, Cleland 2011). This paper is an exercise in philosophical error analysis. I examine this episode with two questions in mind. First, to what extent does the new evidence concerning the colors of dinosaur feathers matter to the debate about the epistemology of historical science that provided the context for the original bet? Second, does this case contain any broader lessons concerning the rationality of placing epistemic bets on the future of natural science? In reply to the first question, I argue that the recent work on the colors of the dinosaurs matters less to the debate about the epistemology of historical science than it might at first seem, in part because there were other problems with the way in which the original debate was framed. In reply to the second question, I argue that this case shows how epistemic bets about the future course of scientific research can go wrong even under nearly ideal conditions. Once we appreciate this, it might seem natural to conclude that we should generally refrain from making such bets. However, it turns out that such a no-betting policy would be too conservative. Figuring out just what sort of policy does make sense turns out to be much more challenging.

Before proceeding, however, one important qualification is in order. One should not come away from this discussion with the impression that trying to figure out the colors of the dinosaurs is representative of the work that paleontologists do. Much work in paleontology does focus on reconstructing prehistoric organisms (for discussion, see, e.g., Turner 2000, Currie 2014 and 2015). But since the 1970s and 1980s, much of the action has focused instead on trying to document and explain large-scale trends and patterns in the fossil record (Turner 2011, Sepkoski 2012). Inferring the colors of the dinosaurs is not too relevant to the big questions about evolutionary patterns and processes that many paleontologists care most about.

2. Recent work on the colors of the dinosaurs. In 2010, an international team of paleontologists from China and the UK published a study of the microstructure of dinosaur feathers from the early Cretaceous period (131-120 mya) from a site in China

(Zhang et al. 2010). Most notably, they looked at feathers from the small therapod dinosaur *Sinosauropteryx*. With the help of a scanning electron microscope, they studied the structure and arrangement of melanosomes, which are tiny pigment-containing organelles found in skin and feathers. Different types of melanosomes are associated with different kinds of pigments. Eumelanosomes contain eumelanin, a black-gray pigment, while phaeomelanosomes contain a pigment called phaeomelanin that ranges in color from reddish brown to yellow. It is possible to look at the shape of melanosomes, as well as the density with which they are packed, and draw some conclusions about coloration. We know this because we know how melanosomes contribute to coloration in the feathers of living birds. And the melanosomes found in the non-avian dinosaur feathers look a lot like the melanosomes associated with the fossilized feathers of ancient birds such as *Confusciornis* (Figure 1). This is a good example of an aspect of historical scientific practice that both Jeffares (2008) and Stanford (2010, 2011) have emphasized: the use of testable regularities to draw conclusions about the deep past. In this case, scientists are taking regularities about how melanosome shape and arrangement are related to color and projecting those back into the past. The fact that the dinosaur melanosomes and the melanosomes in the feathers of modern birds are homologous structures helps to underwrite the inference. This case also affords a good illustration of what Currie (2015) calls the "methodological omnivory" of historical scientists. In their attempts to extract information from the historical record, scientists do not restrict themselves to any one method or inference pattern.

[Figure 1 here.]

Their study of dinosaur melanosomes enabled Zhang and colleagues to make the first ever claim about dinosaur coloration that went beyond educated speculation:

> Only phaeomelanosomes have been identified so far in filaments from the tail of *Sinosauropteryx*, and this suggests that the dark-coloured stripes along the tail in the fossil, and possibly also the filamentous crest along the back, exhibited chestnut to rufous (reddish-brown) tones (Zhang et al. 2010. P. 1077).

Fossil evidence suggests that quite a number of small therapod dinosaurs had feathers or protofeathers. Whether very many ornithiscian dinosaurs had feathers is far less clear, although one recent find suggests that at least some of them did (Godefroit et al. 2014). Not long after Zhang et al. published their groundbreaking study, another team of scientists described in considerable detail the coloration of the late Jurassic therapod *Anchiornis huxleyi*:

> Quantitative comparisons with melanosome shape and density in extant feathers indicate that the body was gray and dark and the face had rufous speckles. The crown was rufous, and the long limb feathers were white with distal black spangles (Li et al. 2010, p. 1369).

They also published a full-color artistic reconstruction of *A. huxleyi* (Figure 2). Thus, in 2010, for the first time, artists' representations of the colors of dinosaurs were based on actual science rather than on speculation and on loose analogies with living organisms. This work on the colors of dinosaur feathers has also lent some support to the view that pennaceous feathers first evolved for purposes of visual display (Koschowitz, Fischer, and Sander 2014). In addition to the original work on *A. huxleyi*, Li and colleagues (2012) reported that the small therapod dinosaur, *Microraptor*, had iridescent black feathers. Carney and colleagues (2012) also reported that the iconic *Archaeopteryx* had black wing feathers, though that finding has since been questioned.

[Figure 2 here.]

In the meantime, seemingly against all odds, scientists have even found some feather filaments from the late Cretaceous preserved in amber (McKellar et al. 2011, Norell 2011). Although some of the feathers look like "dino fuzz" and may have come from non-avian dinosaurs, birds represent another possible source, and the interpretation is somewhat contested (Dove and Straker 2012; see McKellar et al. 2012 for a reply). Even more intriguing than feathers trapped in amber is a recent study that uses trace metals such as copper as biomarkers for particular pigments, including eumelanin (Wogelius et al. 2011). It might be possible to learn something about pigmentation even in cases where melanosome structure is poorly preserved, by looking at the distribution of trace metal biomarkers in a fossilized specimen. Manning et al. (2013) used this approach to challenge Carney's (2012) claim that *Archaeopteryx* had black wing feathers. Manning and colleagues point out that studying melanosome structure only gives us a look at the color at a particular point on the feather. Taking a second look at three *Archaeopteryx* specimens, they used X-ray spectroscopy to do a chemical analysis of a whole feather, with the goal of identifying the distribution of the chemical remnants of eumelanin. They actually found that the animal had lighter colored feathers with black edges and tips. A number of philosophers have emphasized the importance of consilience, or of developing multiple, independent lines of evidence for conclusions about the past (Kosso 2001, Wylie 2002, 2011; Forber and Griffith 2011). We can start to see this happening with the research on the colors of the dinosaurs.

What about dinosaur skin color? A Google search for images of a favorite (probably) featherless dinosaur, such as, say, *Corythosaurus* (so named because its cranial crest resembles a Corinthian soldier's helmet) shows that artists have interpreted their color schemes in various ways: Did the animals have black and white zebra stripes? Or were they dull gray with neon blue cranial crests? It's not clear at this point whether scientists will ever be able to say much about the color of dinosaur skin. So far, just one published study has looked at melanosomes in fossilized skin and scales (Lindgren et al. 2014), and that one focused not on dinosaurs but on three specimens of ancient marine replies: one ichthyosaur, one mosasaur, and one turtle. There are a few dinosaur mummies that include fossilized skin (as opposed to skin impressions), but those are quite rare, as compared with preserved feathers.

Scientists have also begun to study the conditions under which melanosome structure gets preserved. McNamara et al. (2013) ran feathers through an autoclave in order to simulate the fossilization process and see what high temperature and pressure would do to melanosome structure. They found that the melanosomes shrank. Moyer and colleagues (2013) studied the feathers of living chickens in order to develop methods for telling the difference between fossilized melanosomes and microbial impressions. These sorts of studies represent attempts to tease apart just what sort of information gets preserved vs. destroyed in the fossil record.

Needless to say, my prediction that we probably would never know the colors of the dinosaurs turned out to be a bad epistemic bet against science. I made this prediction around the same time that Jakob Vinther, then a graduate student studying the pigmentation in the ink produced by squids, first had the idea to look at melanosome structure (Vinther et al. 2008). In the remainder of this paper, I will explore what these findings about the colors of the dinosaurs might mean philosophically. In section 3, I focus on the epistemological question of how much we can know about the past. Does an episode like this provide grounds for optimism? In section 4, I will consider whether one rational response to this would be to abstain from any such epistemic bets on the future course of science.

3. Implications for the epistemology of historical science. How much can we know about the distant past? Historical processes, such as fossilization, degrade and destroy a great deal of information about the past, a point that Sober (1988, pp. 3-5) made vividly by describing a simple equilibrium model: A ball dropped from the rim of the bowl will eventually come to rest at the bottom, a process that destroys information about the point of release.[1] But history also preserves information, and just how much information

[1] Sober and Steel (2014) explore this idea more formally. They invoke a theorem from information theory (the exponential information loss theorem) to show that where certain conditions are met—where historical processes are both Markovian and regular—information about the past will decay at an exponential rate. They point out that it is an empirical question which, if any, evolutionary processes satisfy those conditions. Their more formal treatment is therefore congenial to the idea that how

is destroyed vs. preserved is ultimately an empirical question (Sober 1988; Jeffares 2010; Tucker 2011). This is, in part, a question about the completeness of the fossil record, and that is itself something that paleontologists study, for example, by experimental replication of some aspects of the fossilization process. Not only that, but some of the central debates of theoretical paleobiology, such as the debate about punctuated equilibria, concern the (in)completeness of the fossil record.

Although it's an empirical question how much information about the past has been preserved, a number of philosophers have at times offered armchair arguments in favor of epistemic optimism or pessimism. My own earlier discussion of the colors of the dinosaurs came in the context of his trying to develop an armchair argument for a more pessimistic view (Turner 2005; 2007, Ch. 2). Around the same time, Kleinhans, Buskes, and de Regt (2005) also argued that underdetermination is especially prevalent in the Earth sciences. By contrast, Cleland (2002, 2011, 2013) has consistently defended a more optimistic view, one that appeals to the time asymmetry of overdetermination, which she takes to be a deep physical fact about our universe. Stanford (2010, 2011) also defends a more optimistic view on the grounds that heterogeneous modes of confirmation are available in historical science. The difference, clearly, is merely one of emphasis. Everyone agrees that we have quite a lot of knowledge of prehistoric life. And presumably, everyone also agrees that there is much that we will never know. But some of us have been tempted by armchair philosophical arguments that have seemed, at least, to support a more optimistic or a more pessimistic view. This temptation is one that we should resist.

---

much information gets preserved vs. destroyed by historical processes is an empirical question.

At first glance, one might think that the recent work on the colors of the dinosaurs lends significant support to epistemic optimism about historical science. How might an argument for optimism go? In my (2005), I treated the colors of the dinosaurs as an example of a local underdetermination problem in historical science, invoking a distinction that Stanford (2001) had drawn between local and global underdetermination. Radical skepticism about the past, of the sort suggested by the question how we can be sure that God did not create the entire universe five minutes ago, is a global underdetermination problem. Those sorts of questions have little to do with actual scientific practice. I characterized local underdetermination problems as any epistemic situations meeting the following conditions:

a. Two incompatible hypotheses, *H* and *H**, are genuine rivals.
b. *H* and *H** are weakly empirically equivalent, meaning that they are equally well supported by all the currently available evidence.
c. As best anyone can tell, *H* and *H** have roughly equal portions of non-empirical theoretical virtue (simplicity, explanatory power, and the like).
d. Background theories give us some reason to think that *H* and *H** are also strongly empirically equivalent (Turner 2005, p. 18).

Two hypotheses are strongly empirically equivalent if and only if they are (or would be) equally well supported by all the empirical evidence that will ever be available (Turner 2005, p. 17). The rough idea is that local underdetermination problems occur when there is an evidential tie between rival hypotheses, and when available background information suggests that the tie will probably not be broken.

I argued that rival hypotheses about the colors of the dinosaurs satisfy conditions (a) through (d) above, and hence count as a local underdetermination problem (2005, 2007, Chapter 2). With hindsight, it may still have been correct at the time to suggest that dinosaur coloration was a local underterdetermination problem. Condition (d) is the important one, and notice how weak it is. Perhaps it was right to say, pre-2008, that background theories about the fossilization process gave us "some reason" to think that we'd never have evidence to discriminate between rival hypotheses about the colors of the dinosaurs. But background theories change. Vinther and colleagues (2008) showed that fossilization processes have the potential to preserve more information than anyone had previously suspected. By relativizing underdetermination to background theories, the above analysis of local underdetermination leaves open the possibility that something that counts as a local underdetermination problem at one time might cease to be so when surprising new evidence comes in, or when background theories get revised.

A different diagnosis of this episode is possible. Perhaps it was a mistake to think, before 2008, that condition (d) was really satisfied. In other words, perhaps it was a mistake, even before the work on dinosaur melanosomes, to say that the colors of the dinosaurs represented a bona fide local underdetermination problem. This second diagnosis is less plausible, however, in part because condition (d) is so weak. Scientists have known about fossilized feathers since the discovery of *Archaeopteryx* in 1861, and prior to Vinther's work, no one suspected that studying the microstructure of those fossils could provide clues about coloration. Although lots of feathered dinosaur specimens were discovered in a set of rock formations known as the Jehol Group in northeastern China beginning in the 1990s, it just never occurred to anyone prior to Vinther to look for melanosomes. The fact that this really was a local underdetermination problem helps explain why the finding was so dramatic. The website of the Paleobiology and Biodiversity Group at the University of Bristol declares that

> The impossible seems to be true: paleontologists can now tell the colours of fossil feathers – in ancient birds and dinosaurs![2]

One of the scientists associated with the Bristol paleobiology group is Michael Benton, a co-author on the paper that first reported the work on melanomes. This bit of public relations is revealing, even if it's a little exaggerated. Before the melanosome work, not only did we have "some reason" to think that information about the colors of the dinosaurs was lost for good, but it seemed very unlikely (even if not impossible) that any relevant evidence would be found.

How might this episode serve as the basis for an argument for epistemic optimism? Here is how the argument might go:

(1) Even in a case where rival hypotheses about the colors of the dinosaurs were locally underdetermined, and where there was ample reason to think that information about dinosaur coloration had been

[2] http://palaeo.gly.bris.ac.uk/melanosomes/, retrieved 30 December, 2014.

destroyed by historical processes, scientists nevertheless developed new methods that enabled them to extract more information from the fossil record.

(2) Therefore, the fossil record probably contains a lot more information than it seems. In most cases, there will be clues—roughly, what Cleland (2002) calls "smoking guns"—waiting to be interpreted that will enable scientists to discriminate between hypotheses that are, for now, locally underdetermined.[3]

The argument has some intuitive force: After all, if paleontologists can solve this problem, which seemed so intractable, that suggests a broadly optimistic answer to the question how much we can know about the past. A couple of philosophers have at least gestured in the direction of this optimistic argument (Cleland 2011, p. 30; Stanford 2010, pp. 238-9). But this argument is one that we should resist. If a few examples of underdetermination do not support a generally pessimistic view of the epistemology of historical science, then neither do examples of success support a more sanguine view. Indeed, the optimistic argument as formulated above looks like one more attempt to defend epistemic optimism (or pessimism) from the armchair.

[3] Forber and Griffith (2011, p. 3) read Cleland as saying that a smoking gun is the rough equivalent, in historical science, of a crucial experiment. It is probably more charitable to read her as saying that a smoking gun is just a trace (or set of traces) that confirms one historical hypothesis over some rival. On this wider reading, which is shared by Currie (forthcoming), fossilized melanosomes would count as a smoking gun for dinosaur coloration.

One problem with the argument for optimism is just that a single case of epistemic success, however, dramatic, does not support any general optimistic conclusions. One need not look too far to find other good examples of local underdetermination problems in historical science, and the successful inference of dinosaur coloration must be balanced against those others. In their recent book of paleo art, Conway, Kosemen, and Naish (2013) drive this point home vividly with their nonstandard depictions of dinosaurs. They portray dinosaurs in ways that are completely different from what we are used to seeing in museum exhibits, and yet perfectly consistent with the available fossil evidence. For example, they question the usual tendency to portray dinosaurs as especially lean, and they paint a picture of a chubby *Parasaurolophus* (p. 51). To drive the point home, they show a bizarre streamlined cow, drawn in the style that paleo artists use with dinosaurs. The successful inference of coloration in some dinosaurs does not imply that underdetermination is not a pervasive problem.

A second problem with the argument for optimism is that whenever scientists settle one question in dramatic fashion, new, often finer-grained questions crop up, and the answers to those new questions may well be underdetermined. For example, the fossil record also exhibits what is sometimes called developmental incompleteness. For some taxa, we may have no good fossil remains of juveniles, or else it may be difficult to tell whether two specimens represent different species or individuals of the same species at different life stages. What color were the hatchlings of *Sinosauropteryx*? Did their feather coloration change over the course of development? Maybe we'll eventually be able to answer these questions, and maybe we won't. Once we get empirical traction with respect to something we thought was an underdetermination problem, new finer-grained questions usually come into view, and the answers to those may well be locally underdetermined.

Forber and Griffith (2011) make a similar point in the course of developing their response to Cleland's (2002) discussion of the Alvarez impact hypothesis. We now have multiple lines of evidence pointing toward the conclusion that an asteroid hit the earth at the end of the Cretaceous period. But that raises new, finer-grained questions about how to explain the distinctive pattern and selectivity of extinctions that occurred. Complicating matters is the fact that some lineages disappear from the fossil record before the K-Pg boundary. Might other geological processes, such as the volcanism that produced India's Deccan traps, have played some contributing role? Forber and Griffith write that

> while there is little doubt that an impact played some role in these extinctions, the available data currently fail to discriminate between the strong impact hypothesis and that of multiple, smaller, temporally proximate extinction events, or those involving a requisite role for other geological processes (p. 13).

The confirmation of the impact hypothesis was a major epistemic success, but every such success raises new questions.

To return to the main thread, consider that modern birds are tetrachromats, meaning that they have four different kinds of cone cells.[4] Humans and other primates,

[4] I thank Olivia Ziegler for calling my attention to this issue.

by contrast, are trichromats, while most other mammals are dichromats. Birds can perceive a much wider range of colors than we can, including parts of the ultraviolet spectrum. It appears that tetrachromacy was the ancestral trait, and that the vision system in mammals became simplified early in their evolutionary history, probably during a long stretch of nocturnal living (Bowmaker 1998, 2008). Phylogenetic reasoning suggests that non-avian dinosaurs were probably tretrachromats, just like modern birds and reptiles (Rowe 2000; Koschowitz, Fischer, and Sander 2014). If this is correct, then it could be difficult to determine what the dinosaurs would have looked like to each other; at best, the recent scientific work is telling us what they would have looked like to trichromats like us. This is not merely a philosophical puzzle about other minds in the sense of Nagel (1974). Scientists who study color perception in modern birds have demonstrated experimentally that many birds respond to ultraviolet cues in contexts such as foraging and mate selection. (For a review of some of this research, see Bennett and Théry 2007.) Insofar as we care about the evolution of coloration in dinosaurs, how their plumage would have looked to other dinosaurs really matters. How they would have looked to us trichromats is now easier to answer, but arguably less important.

To draw some intermediate conclusions: First, it was probably correct to say, pre-2008, that dinosaur coloration was an example of a local underdetermination problem. Second, the dramatic success in this case does not support any generalized optimism about the prospects for answering questions about the deep past, in part because epistemic successes often point toward finer-grained underdetermination problems. Third, the question how much we can know about the past is an empirical one, not one that can be settled by armchair philosophical arguments. The question whether we should be generally optimistic or pessimistic when it comes to how much we can know about the deep past is perhaps not a terribly fruitful one to begin with.

4. An argument against epistemic betting. One response to my failed epistemic bet against dinosaur science would be to adopt a policy of abstaining from any further predictions about which questions scientists will or will not be able to answer in the future. This policy of abstention from epistemic betting seems like the obvious lesson to draw from this case. A "no-betting" policy might receive some additional motivation from reflecting on other past cases where people predicted that science would never be able to figure something out. For example, in his *Critique of Judgment*, Immanuel Kant famously predicted that there will never be a "Newton of the blade of grass":

> [w]e can say boldly it is alike certain that it is absurd for men . . . to hope that another *Newton* will arise in the future, who shall make comprehensible by us the production of a blade of grass according to natural laws which no design has ordered (Kant 1790/1914 §75, pp. 312-13)

Just how to read this famous passage is a matter of scholarly discussion, but it seems fair to interpret Kant as saying that biologists will never be able to give a satisfactory causal-mechanical explanation of organisms, a prediction that turned out to be mistaken, thanks to Darwin (for discussion, see Cornell 1986).

The Kant example suggests that a healthy intellectual humility about science might help to motivate a "no betting" policy. Science is full of surprises, and many episodes from the history of science involve difficult-to-foresee twists and turns. That history also exhibits quite a lot of contingency, where seemingly small developments or methodological innovations can make a big difference to subsequent work. At first blush, it seems like a "no-betting" policy would reflect appropriate epistemic caution.

An argument in favor of the no-betting policy might run as follows:

(1) Pre-2008, our understanding of the fossilization process did give us some reason to think that we would not find any clear evidence of the colors of the dinossurs. At that time, it would have been reasonable to predict that scientists would not be able to infer dinosaur coloration.

(2) Yet scientists did find ways to test claims about the colors of the dinosaurs.

(3) Therefore, even the most reasonable bets on the future course of historical scientific work are liable to turn out wrong.

(4) In light of (3), caution dictates that one should refrain from betting on the future course of historical scientific research.

At first glance, this seems like quite a compelling argument. Some of the considerations of the previous section support premise (1). Someone who has doubts about premise (1) might consider the epistemic situation of someone at an earlier time—say, in the 1980s, before the discovery of so many Chinese feathered dinosaurs. Surely it would have been reasonable then to conclude that we probably would not find evidence of the colors of the dinosaurs. That bet, made at an earlier time, would also have been a losing one. Here I want to focus primarily on the move from (3) to (4). Does the cautious conclusion (4) really follow from (3)? For present purposes, I framed (4) rather narrowly, as pertaining only to historical science. One might think, however, that the lesson is more general than that: Apparently reasonable bets about which questions scientists will or will not be able to answer in the future are liable to come out wrong. So we should adopt a policy of not betting on science at all. I want to suggest that this no-betting argument is too conservative, for two reasons. The main reason for adopting a "no betting policy" is epistemic caution, but there is such a thing as being too cautious.

First, if we accepted conclusion (4), then we would never be able to identify any local underdetermination problems at all. That is because identifying an underdetermination problem involves making a kind of epistemic bet. But this result is just implausible. Of course we can safely identify some local underdetermination problems. We will probably never know, for example, the exact size of the Tyrannosuar population immediately prior to its extinction. If we were to bet, now, that scientists will never determine the size of the Tyrannosaur population just before the asteroid hit, that would be just another bet against future paleontology, perhaps a bit safer than my bet concerning the colors of the dinosaurs, but no different in kind. Thus, the main problem with the cautious conclusion (4) is just that some epistemic bets on the future course of science still seem perfectly reasonable. We make such bets any time we identify a local underdetermination problem.

A second problem with the cautious conclusion (4) is that scientists themselves place such bets all the time. Indeed, the identification of local underdetermination problems is an important though underappreciated feature of historical scientific practice. Perhaps part of the neglect is due to the fact that this issue seems to belong more to the context of discovery as opposed to the context of justification, whereas most philosophers of science are interested in normative questions about evidence and confirmation. But there are normative (as opposed to psychological and historical) questions in play here as well. At the very least, the case of the colors of the dinosaurs raises the question under what conditions it's rational to place epistemic bets on the future course of science. Most of the recent work done by philosophers of historical science has focused on the ways in which scientists confirm or disconfirm claims about the past. But scientists also routinely draw inferences about the future of their own fields.

Historical scientists and the institutions that fund their work have to make decisions about which questions are worth pursuing and which are best left for another day, or bypassed completely. Such decisions are an important part of scientific practice, and they surely involve complicated pragmatic cost/benefit calculations. They also involve epistemic betting of precisely the sort that (4) would counsel against. One challenge is to identify questions that seem answerable using available methods, and where there is some reasonable expectation of finding or producing relevant evidence. To revert to an earlier example, there's no point in devoting time or resources to trying to determine the exact size of the Tyrannosaur population just prior to its extinction. For another example, we now have good methods for investigating the colors of dinosaur feathers, but not the colors of their eyes. Eye color is more complicated because it involves not only pigmentation of the iris (which isn't typically preserved in the fossil record) but also the scattering of light as it passes through the iris. Deciding which problems to work on involves epistemic betting, and so would run afoul of the conservative conclusion (4).

One potential worry here is that the cases where it seems rational to bet against science might also seem scientifically uninteresting. Who cares about the exact size of the Tyrannosaur population at the moment the steroid hit? One response to that worry to point to ways in which knowing the *T. rex* population size at a given time might prove useful: For example, it might matter if we were interested in predator-prey ratios, or in the relationship between abundance and extinction risk. Second, we should be mindful of a circularity problem here. It's plausible that our epistemic resources inform our judgments about what counts as interesting. In cases where we know we have no scientific tools that give us any traction we might be more likely to dismiss questions as trivial or uninteresting. On the other hand, the fact that we do have tools that give us some empirical traction with respect to some question can make that question seem interesting and important, if only because it affords us an opportunity to put our epistemic tools to work. Interestingness therefore may be partially relative to our epistemic resources. Today, part of what makes the colors of the dinosaurs interesting is just that scientists have developed tools that give us epistemic access. If we had some way of ascertaining the size of the Tyrannosaur population when the asteroid hit, then that issue might come to seem more interesting.

The above reflections on interestingness point to another important feature of this case. If I am right, scientists often make epistemic bets about future evidence; they identify local undertermination problems—like the size of the *T. rex* population when the asteroid hit—only to set them to one side indefinitely. But the popular portrayal of full color dinosaurs in books and films makes it difficult to set the issue to one side in quite the same way. This made dinosaur coloration especially interesting as a potential underdetermination case: it seems important that popular reconstructions were, and mostly still are, so speculative. But this may also have contributed to the riskiness of my bet against dinosaur science. It would be much safer to bet against the science in cases where there is less pressure on scientists to generate answers, and less reward for those who do.

Even if there is nothing wrong with epistemic betting, *per se*, it might still be possible to identify other features of my bet against dinosaur science that made it especially problematic. The conservative no-betting argument clearly goes too far, but perhaps there are certain types of epistemic bets that we should eschew. One possible source of trouble for my earlier treatment of the colors of the dinosaurs as an underdetermination problem was the time frame involved. Consider the difference between saying that scientists will probably never have the evidence they need to settle some particular question, and saying that we're not likely to get the evidence we need anytime soon. Claiming that we'll probably never figure something out goes a lot further than one would ever need to in the context of scientific practice. The weaker claim that we probably won't get the evidence we need anytime soon would be enough to motivate a decision to work on something else. One plausible suggestion, then, is that we should think of the time frame as a variable that corresponds with the riskiness of an epistemic bet, while refraining from medium to longer-term bets. Perhaps the relatively safer, shorter term bets are all that's needed for scientific practice. Note that condition

(d) in the above analysis of local underdetermination problems also makes no reference to time frame. If we wanted to, we could adjust that condition so as to include some reference to the time frame of the prediction. Abstaining from all epistemic betting is too conservative, but it might be rational to adopt a policy of avoiding relatively riskier, longer-term bets.

As plausible as it seems, a policy of abstaining from relatively longer-term bets still runs into problems. First, we have already seen that there are clear-cut cases of perfectly safe long-term epistemic bets. Think again of the clear-cut underdetermination cases, such as the size of the Tyrannosaur population when the asteroid hit. In the face of such cases, the proposed policy still seems too conservative. A more serious problem, however, is that it would have permitted a short-term bet against scientists' figuring out the colors of the dinosaurs. In 2005, even a short-term bet would have come out wrong. So it's hard to see why a case where even a short-term bet would have come out wrong should motivate a moderately conservative policy of abstaining from relatively longer-term bets. In this particular case, the time frame was not the source of the problem.

One plausible suggestion is that the problem with my bet against dinosaur science was that it failed to take into consideration the fact that scientists are often working to improve their methods and techniques.[5] Consider the difference between:

B1. Betting that we will probably never be able to determine the colors of the dinosaurs using available, familiar methods and techniques.

[5] I thank Adrian Currie for suggesting the approach developed in this paragraph.

and

> B2. Betting that future improvements and refinements in our methods and techniques will probably not enable us to determine the colors of the dinosaurs.

B1 would have been completely reasonable pre-2008. Call that a *current methods bet*. However, B2 might seem to have gone too far, given the difficulty of predicting future methodological improvements. Call B2 a *methods neutral bet*. One alternative diagnosis of my bet, then, was that it was ambiguous between B1 and B2. This suggests a different sort of refinement of the conclusion of the no betting argument: perhaps the argument only supports a policy of abstaining from methods neutral bets.

This moderately cautious conclusion is extremely plausible, in part because (with hindsight) it would have provided the right sort of guidance with respect to the colors of the dinosaurs. If one had followed this moderately conservative policy pre-2008, then one could have made the perfectly safe current methods bet. But *that* bet actually would have come out right. It remains the case that no one has been able to make out the colors of the dinosaurs using pre-2008 methods. Note also that condition (d) in the analysis of local underdetermination is methods neutral. If background theories give us some reason to think that two rival hypotheses are strongly empirically equivalent, then those background theories are giving us some reason to think that even future methodological innovations will not enable us to discriminate between the two hypotheses.

As plausible as it is, a policy of abstaining from methods neutral betting may still be too conservative. There are cases—such as the exact size of the global Tyrannosaur population at the moment when the asteroid hit, or whether it rained on this spot 3,000 years ago—where even methods neutral bets seem perfectly safe. Typically, these cases of apparently safe methods neutral bets will involve fine-grained historical particulars. And scientists typically do not waste their time investigating them. These cases seem trivial, but (as noted above) our judgment that they are is likely informed by methods neutral betting.

This last thought suggests that methods neutral epistemic bets may yet have an important role to play in scientific practice. It's unproblematic to say that current methods probably will not enable anyone to determine the exact size of the Tyrannosaur population just prior to extinction. Current methods provide no traction here. But the decision not to waste one's time trying to find some new way to answer a question like this one would have to be based on *more* than a current methods bet. If the current methods bet were all one permitted oneself to make one could still waste time trying to develop new methods to tackle this question. That would probably be a waste of time because of what we already know about how fossilization works. Since only a tiny percentage of the members of any given species end up in the fossil record, the record just doesn't contain information about the size of a given population at a particular moment in history. Cases like this suggest that methods neutral betting does matter to scientific practice. So prohibiting it might be a mistake.

Another reason not to prohibit methods neutral bets is that there are other contexts, less central to the everyday practice of science, where one might want to make them. Consider the prediction that future methodological innovations will never enable scientists to determine whether there is an afterlife. This is a good example of what I am calling a methods neutral bet, and it's one that many scientists and philosophers would be comfortable with. Indeed, such a bet on the future course of science might help to motivate methodological naturalism, which rules hypotheses about the supernatural out of bounds for purposes of scientific investigation. These issues deserve more attention than I can give them here (but see Sober [2011] for a wide-ranging discussion of methodological naturalism). I raise the issue of methodological naturalism only to suggest that a blanket prohibition against methods neutral betting might go too far, because it would potentially rule out a number of epistemic bets on the future of science that we might want and need to make. If we prohibited *all* methods neutral bets, we'd have to abstain from predicting that science will probably never be able to tell us about the afterlife.

Questions about the rationality of placing bets on the future of science remain under-explored. I've tried here to make some first steps toward thinking through the issues, but complications loom. For example, I have been arguing that epistemic betting, including even methods neutral betting, is an important feature of scientific practice. If this is right, then pessimistic predictions made by scientists could be self-fulfilling. Consider a situation in which the scientific community agrees that we probably won't ever be able to discriminate between two rival hypotheses, *H1* and *H2*. So they set the issue to one side indefinitely. With no one working on developing new methods to tackle the problem, it becomes a foregone conclusion that we'll never be able to discriminate between *H1* and *H2*. But suppose that if scientists had not set the issue to one side, someone like Jakob Vinther would have surprised everyone by developing new empirical methods that help decide the issue. In such a scenario, the scientific community's epistemic betting would be self-fulfilling. That is, we'd have a case of local underdetermination that scientists sustain indefinitely by means of their own epistemic betting—a case of manufactured undeterdetermination.

To sum up the arguments of this section: A no-betting policy would be too conservative, and focusing on the time frame (i.e. relatively shorter- vs. longer-term) misdiagnoses the problem with my bet against dinosaur science. The distinction between current methods bets (which seem relatively unproblematic) and methods neutral bets captures something important, but a blanket prohibition against methods neutral bets still seems too conservative, mainly because we can still point to clear examples of safe methods neutral bets. But also because even methods neutral bets plausibly have a role to play in scientific practice. More generally, my failed bet against dinosaur science highlights an issue about which we just don't (yet) have much good philosophical theory: under what conditions is it rational to make claims about what questions scientists will or won't be able to answer in the future? The analysis of this episode also suggests a possibility that philosophers have not considered—namely, that scientists' own identification of local underdetermination problems might be self-fulfilling.

5. Conclusion. Figuring out the colors of the dinosaurs is somewhat peripheral to paleontologists' efforts to reconstruct the big picture of evolutionary history. Nevertheless, my mistaken prediction that scientists would not be able to figure out the colors of the dinosaurs is worth revisiting. For one thing, it is important to be clear about where the real mistake occurred. It was probably right to treat the colors of the dinosaurs as a local underdetermination problem. The mistake was in thinking that particular examples of underdetermination support a generally pessimistic view of historical science. Anyone who draws an optimistic conclusion from this example of scientific success commits the same error, but in the other direction. This episode matters less to the epistemology of historical science than it might seem.

In another way, however, the case of the colors of the dinosaurs matters more than it might seem. For this episode also raises some general questions about the rationality of placing epistemic bets concerning which questions science will be able to answer in the future. In this paper, I have tried only to develop some opening arguments concerning an issue that philosophers have not focused on much. I've argued that a blanket policy of abstaining from such bets goes to far. I have also entertained, and ultimately dismissed, policies of abstaining from certain kinds of bets—long-term ones, as well as methods neutral bets. Given the importance of epistemic betting to the actual practice of science, and the possibility that such bets might be self-fulfilling, the question of which policy would be most rational is one that could use more sustained philosophical attention.

Acknowledgments

I have received a lot of help with this paper, and am grateful to many friends and colleagues, especially Adrian Currie, Simon Feldman, Greg Radick, Peter Vickers, and Alison Wylie. Earlier versions of the paper were presented at Durham University, UK, in September 2012, and at the workshop on Rock, Bone, and Ruins: Evidence and Confirmation in Historical Science, University of Sydney, Centre for the Foundations of Science, May 2014. I'm grateful to audiences at both of those events for helpful feedback. I also received good suggestions from two anonymous reviewers for this journal.

References

Bennett, A.T.D. and M. Théry. 2007. "Avian color vision and coloration: Multidisciplinary evolutionary biology," *American Naturalist* 169: S1-S6.

Bowmaker, J.K. 1998. "Evolution of colour vision in vertebrates," *Eye* 12: 541-547.

Bowmaker, J.K. 2008. "Evolution of vertebrate visual pigments," *Vision Research* 48: 2022-2041.

Carney, R., et al. 2012. "New evidence on the colour and nature of the isolated Archaeopteryx feather," *Nature Communications* 3: 637.

Cleland, C. 2002. "Methodological and epistemic differences between historical science and experimental science," *Philosophy of Science* 69(3): 474-496.

Cleland, C. 2011. "Prediction and explanation in historical natural science," *British Journal for Philosophy of Science* 62(3): 1-32.

Cleland, C. 2013. "Common cause explanation and the search for a smoking gun," in V. Baker, ed., *Rethinking the Fabric of Geology*, Special Paper 502. Geological Society of America, pp. 1-9.

Conway, J., C.M. Kosemen, and D. Naish. 2013. *All Yesterdays: Unique and Speculative Views of Dinosaurs and Other Prehistoric Animals.* Lexington, KY: Irregular Books.

Cornell, J.F. 1986. "Newton of the Grassblade? Darwin and the Problem of Organic Teleology," *Isis* 77(3): 404-421.

Currie, A. (forthcoming), "Hot-Blooded Gluttons: Dependency, Coherency, and Method in the Historical Sciences," *British Journal for the Philosophy of Science*.

Currie, A. (2015) "Marsupial lions and metholodological omnivory: Function, success, and reconstruction in paleobiology," *Biology and Philosophy*.

Currie, A. (2014), "Narratives, mechanisms, and progress in historical science," *Synthese* 191(6): 1163 – 1183.

Dove, C.J., and L.C. Straker, 2012, "Comment on 'A diverse assemblage of late Cretaceous dinosaur and bird feathers from Canadian amber'," *Science* 335: 796.

Forber, P., and E. Griffith, 2011, "Historical Reconstruction: Gaining Epistemic Access to the Deep Past," *Philosophy and Theory in Biology*, available online at http://quod.lib.umich.edu/p/ptb/6959004.0003.003?rgn=main;view=fulltext, retrieved 15 June, 2015.

Godefroit, P., et al. 2014. "A Jurassic ornithiscian dinosaur from Siberia with both feathers and scales," *Science* 345(6195): 451-5.

Jeffares, B. 2008, "Testing times: Regularities in the historical sciences," *Studies in History and Philosophy of Biology and Biomedical Sciences* 39: 469-475.

Jeffares, B. 2010. "Guessing the future of the past," *Biology and Philosophy* 25(1): 125-142.

Kant, I. 1790/1914. *Critique of Judgment*, second edition. J.H. Bernard, trans. London: Macmillan and Co.

Kemp, T.S. 1999. *Fossils and Evolution*. Oxford: Oxford University Press.

Kleinhans, M.G., Buskes, C.J.J., and H. de Regt. 2005. "Terra Incognita: Explanation and Reduction in Earth Science," *International Studies in the Philosophy of Science* 19(3): 289-317.

Koschowitz, M.-C., C. Fischer, and M. Sander. 2014. "Beyond the rainbow," *Science* 346(6208): 416-7.

Kosso, P. 2001. *Knowing the Past: Philosophical Issues of History and Archaeology*. Amherst, NY: Humanity Books.

Li, Q., et al. 2010. "Plumage color patterns of an extinct dinosaur," *Science* 327(5971): 1369-72.

Li, Q., et al. 2012. "Reconstruction of *Microraptor* and the evolution of iridescent plumage," *Science* 335: 1215-1219.

Li, Q., et al. 2014. "Melanosome evolution indicates a key physiological shift within feathered dinosaurs," *Nature* 507: 350-3.

Lindgren, J. et al. 2014. "Skin pigmentation provides evidence of convergent melanism in extinct marine reptiles," *Nature*, 506: 484-488.

Manning, P.L., et al. 2013. "Synchrotron-based chemical imaging reveals plumage patterns in a 150 million year old early bird," *Journal of Analytical Atomic Spectrometry* 28: 1024-1030.

McKellar, R.C., Chatterton, B.D.E., Wolfe, A.P., and P.J. Currie, 2011, "A diverse assemblage of late Cretaceous dinosaur and bird feathers from Canadian amber," *Science* 333: 1619-22.

McKellar, R.C., Chatterton, B.D.E., Wolfe, A.P., and P.J. Currie, 2012, " Response to Comment on 'A diverse assemblage of late Cretaceous dinosaur and bird feathers from Canadian amber'," *Science* 335: 796.

McNamara, M.E., et al. 2013, "Experimental maturation of feathers: implications for reconstructions of fossil feather colour," *Biology Letters* 9: 20130184.

Moyer, A.E., et al. 2013, "Melanosomes or microbes: Testing an alternative hypothesis for the origin of microbodies in fossil feathers," *Scientific Reports* 4, doi:10.1038/srep04233

Nagel, T. 1974, "What is it like to be a bat?" *Philosophical Review* 83: 435-450.

Norell, M. 2011, "Fossilized feathers," *Science* 333: 1590-91.

Rowe, M.P. 2000, "Inferring the retinal anatomy and visual capacities of extinct vertebrates," *Palaeontologia Electronica* 3(1). Available online at http://palaeo-electronica.org/2000_1/retinal/text.pdf, retrieved 21 June, 2015.

Sepkoski, D. 2012. *Rereading the fossil record: The growth of paleobiology as an evolutionary discipline*. Chicago, IL: University of Chicago Press.

Sober, E. 1988. *Reconstructing the Past: Parsimony, Evolution, and Inference*. Cambridge, MA: MIT Press.

Sober, E. 2011. *Did Darwin write the* Origin *backwards? Philosophical essays on Darwin's theory*. Amherst, NY: Prometheus Books.

Sober, E., and M. Steel. 2014. "Time and knowability in evolutionary processes," *Philosophy of Science* 81(4): 558-579.

Stanford, P.K. 2001, "Refusing the Devil's bargain: What kind of underdetermination should we take seriously?" *Philosophy of Science* 68(Supplement): S1-S12.

Stanford, P.K. 2010. "Getting real: The hypothesis of organic fossil origins," *The Modern Schoolman* LXXXVII: 219-241.

Stanford, P.K. 2011. "Damn the consequences: Projective evidence and the heterogeneity of scientific confirmation," *Philosophy of Science* 78(5): 887-899.

Tucker, A. 2011. "Historical science over- and under-determined: A study of Darwin's inference of origins," *British Journal for the Philosophy of Science* 62(4): 805-829.

Turner, D. 2000. "The Functions of Fossils: Inference and Explanation in Functional Morphology," *Studies in History and Philosophy of Science C: Biology and Biomedical Sciences* 31: 193-212.

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.

Vinther, J., D.E.G. Briggs, R.O. Prum, and V. Saranathan. 2008. “The colour of fossil feathers,” *Biology Letters* 4: 522-525.

Wylie, A. 2002. *Thinking from Things: Essays in the Philosophy ofArchaeology*. Berkeley, CA: University of California Press.

Wylie, A. 2011. “Critical distance: Stabilizing evidential claims in archaeology,” in P. Dawid, W. Twining, and M. Vasilaki (eds.), *Evidence, Inference, and Inquiry*, Oxford University Press, pp. 371-394.

Zhang, F., et al. 2010. “Fossilized melanosomes and the colour of Cretaceous dinosaurs and birds,” *Nature* 463: 1075-8.

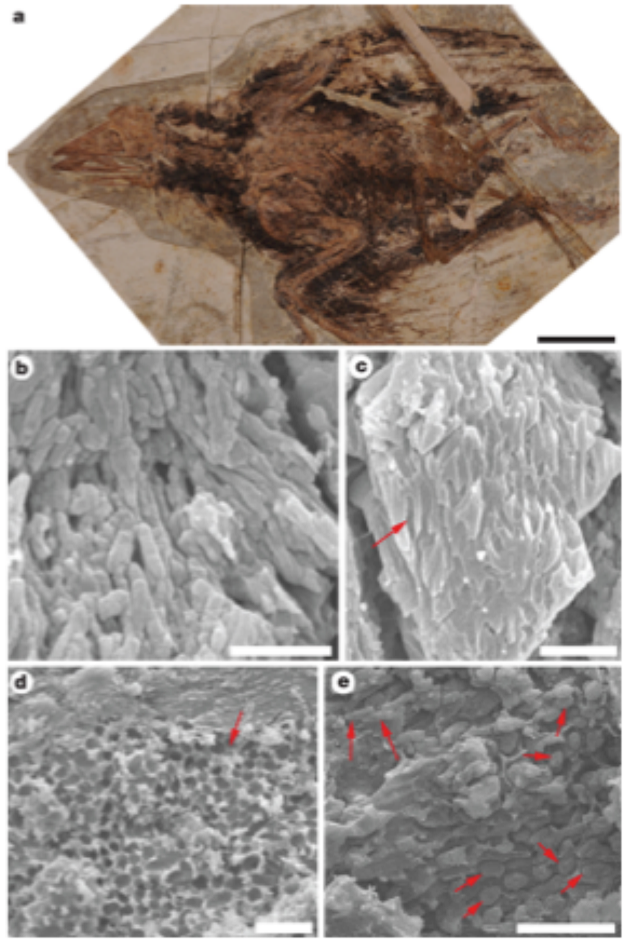

Figure 1. Fossilized melanosomes from the ancient bird, *Confusciornis*. Reprinted by permission from Macmillan Publishers Ltd.: Zhang, F., et al. 2010. "Fossilized melanosomes and the colour of Cretaceous dinosaurs and birds," *Nature* 463: 1075-8. The oblong structures in b and c are fossilized phaeomelanosomes. d and e show a mixture of oblong phaeomelanosomes and more spherical eumelanosomes. In e, the longer arrows indicate phaeomelanosomes, while the shorter arrows point to eumelanosomes.

Figure 2. Artistic rendering of *Anchiornis huxleyi*. The crown is depicted as reddish-brown. The leg feathers are white with "black spangles." From Li, Q., et al. 2010. "Plumage color patterns of an extinct dinosaur," *Science* 327(5971): 1369-72. Reprinted with permission from AAAS.