Introduction

Many of science's most revolutionary discoveries concern *deep time*: the past stretching beyond memory and written texts. One discovery was deep time itself. In Europe at least, the realization that the Earth outruns by millennia both biblical history and our own species' existence shook conceptions of humanity's place in the world just as surely as earlier astronomical discoveries (Rudwick, 2014). But the past isn't simply long. Before Homo sapiens, the climate and continents shifted while diverse lineages arose, became extinct and others evolved. We are the result of millions of years of evolution, a heritage that shapes and constrains how we adapt to our still-changing world. Before written records, then, there wasn't simply a past – there was *history.* Scientific understanding of the deep past emerged in the nineteenth century and crucial aspects, how plate tectonics shape geography and climate for instance, have only been accepted in the last half-century (Oreskes, 1999). Considering the extent to which extinction, evolution, plate tectonics and deep time itself form the furniture of our conceptions of the world and our place within it, their recent pedigree is startling.

In what follows, we'll examine the nature and epistemology of this 'deep past'; what we might call 'prehistorical history'. How do historical scientists reach beyond human memory? How does the nature of the past constrain our knowledge? How does history matter for knowing?

Our central question, then, concerns the relationship between history and knowledge. In making inferences, as well as understanding and explaining the world, does history matter? There are at least two ways in which it might. First, is there something special about trying to understand processes or entities located in the past, as opposed to in the present or future? That is, are there systematic claims to be made about the epistemic status of past things in virtue of their being past? Second, does something's history matter to the knowledge we can have of it? The former question concerns whether a target's being in the past makes it special qua object of knowledge. The latter concerns whether a target's history matters for how we might come to know it.

Regarding that first question, I'll draw a negative conclusion: processes and patterns in the past do not represent a fundamentally different kind of epistemic target from those occupying different temporal locations. Regarding the second question, I'll draw a positive conclusion, for two reasons: (1) the past matters for all scientific inference because the provenance of an inference's data always matters; (2) history matters because how a target came to be can make a difference to what knowledge we can have of it. Note an ambiguity: by 'history' do we mean simply being past, or do we mean something richer – having a particular kind of past, or instantiating a particular kind of dynamics?

I'll typically reserve the term 'history' for this latter notion: 'history' doesn't so much pick out a temporal location (the past) as it picks out a set of events, entities or processes for whom distinctive, 'historical' features make a difference to how we might know them. What might these historical features be? Well, that question is in a nutshell what this Element is about.

We'll also ask after the epistemic value of such enquiries. I'll suggest that investigating the deep past involves more than considering particular histories: we delve into the great diversity of forms, structures, trajectories, events, entities and processes that constitute and shape the world, and the conditions enabling them; we bring contemporary conceptions to the past in analysis, data-gathering and interpretation, and that past in turn shapes those conceptions. Thus, history matters for knowledge, and the process of understanding the deep past is rich and invaluable.

Finally, knowing about the past is necessary for understanding much of what we'd like to know. We occupy an often unrepresentative, atypical sliver of time. Our immediately accessible sample is biased, extremely incomplete, inadequate to answer questions at long scales (Marshall, 2017). These questions are Big. How does evolution work? How do planets, solar systems and the universe form? What explains geographical patterns: mountain ranges, valleys? How did our species evolve and radiate across the globe? And these questions matter: How do species become extinct, how do changes in atmospheric composition alter global temperatures and how do changes in global temperatures affect everything else? Answering these questions requires evidence and perspectives that overcome the inherent bias of our little sliver: a long-term view into the deep past. History matters at least because knowledge of it is necessary for answering Big Questions.

This Element is divided into three related parts. In the first, I consider the relationship between evidence, justification and the deep past. In the second, I consider the nature and contingency of history. In the third, I turn to narrative explanation and its role in history. I conclude with a discussion of the value and purpose of historical science itself. I haven't quite written an introduction to the philosophy of the historical sciences. It is instead an extended essay on the relationship between history and knowledge. Given the aims of the Elements series – to be accessible yet substantive – I've aimed for ambition over completeness. Better, I think, to push the boat into deep waters and risk foundering than to stick to cautious shallows.

1 History and Evidence

History and evidence are intertwined: to ask certain questions we need a long-term scale; to uncover the past we require at least some of its remnants to join us in the present. But is there something special about historical evidence: is it particularly difficult, or impoverished or privileged? Some have thought so.

Recently, Derek Turner (2005, 2007) argued that historical evidence is systematically less powerful than experimental evidence; Carol Cleland (2002, 2011, 2013) argued that historical evidence underwrites a distinctive method that is at least equal to more familiar forms of scientific knowledge. Such arguments appeal to some fundamental ontological or epistemological differences between the past on the one hand, and the present and future on the other. For Turner, investigating the past denies us the boon of repeated experimental investigation; for Cleland, investigation of the past grants the boon of bountiful traces.

I'll argue that although in historical contexts evidence's past matters, this is true of all evidence, and so carries with it no special insight about historical knowledge per se. There is nothing distinctive, epistemically speaking, about past objects of knowledge. I'll begin with a methodological discussion, arguing that to understand how history matters for knowledge, we should begin by understanding the practices that generate such knowledge. This motivates examining historical reasoning 'in play': we'll look at recent work on dinosaur development. Based on that case study, I'll then characterize 'trace-based reasoning', presenting and resolving a puzzle concerning it, before considering the relationship between experimental reasoning and trace-based reasoning. I'll conclude that no evidential reasoning escapes history; however, a target's being in the past doesn't in and of itself raise distinctive epistemic challenges.

1.1 The Very Possibility of Historical Knowledge

Where should our project begin? That is, to understand the relationship between knowledge and the deep past, which philosophical approach is appropriate? I think our starting point should be the practices of historical scientists themselves, but let's consider a few options to see why.

Perhaps we should ask after historical knowledge's very possibility: what is necessary for justified knowledge about any past occurrence? Roughly a century ago, Bertrand Russell posed a thought experiment (Russell, 1921). Imagine that the world blinked into existence five minutes ago: a world which is in every way identical to the world as it is, except its past is only 300 seconds long. An identical duplicate, containing all the same memories, fossils, elementary particles and so forth, as our world. Can we tell whether we live in a billions-of-years-old or 300-seconds-old universe? As Russell points out,

> There is no logical impossibility in the hypothesis that the world sprang into existence five minutes ago, exactly as it then was, with a population that ‘remembered’ a wholly unreal past. There is no logically necessary connection between events at different times; therefore nothing that is happening now or will happen in the future can disprove the hypothesis that the world began five minutes ago (Russell, 1921, p. 158).

All present evidence *underdetermines* the hypothesis that the world has a long past and the hypothesis that the world blinked into existence 5 minutes (or 20 minutes, or 3 seconds) ago. ‘Underdetermination’ is a relationship between at least two hypotheses and a body of evidence: when the evidence is insufficient to decide between those hypotheses, they are underdetermined by it (Wylie, 2019; Godfrey-Smith, 2008; Laudan, 1990). Under Russell’s scenario any observations we might make now are the same whether or not the past existed for billions of years or 300 seconds. However, under the sceptical hypothesis, all claims about the past beyond those five minutes appear to come out false, while under the standard hypothesis some at least are true. In one world it is approximately true that (non-avian) dinosaurs died out 65.5 million years ago, in the other, (ignoring philosophical finagling about the nature of truth and reference) it is false.

What are we to make of such hypotheses? We might take our cue from Russell himself: ‘I am not here suggesting that the non-existence of the past should be entertained as a serious hypothesis. Like all sceptical hypotheses, it is logically tenable but uninteresting’ (Russell, 1921, pp. 159–160).

I don’t think Russell is quite right about this.[1] One take-home message is that our knowing about the past depends on features of the past itself: in this extreme case it depends on the past existing (or having had existed). If our knowledge of the past is to be like other empirical knowledge, then it is predicated on there being something that our knowledge is about. Russell is right that, as a serious hypothesis, the five-minute-hypothesis is rather uninteresting. After all, if no possible observation might make a difference to our status as knowers, what are we to do but shrug? But the thought that knowledge of the past depends in part on its nature deserves reflection (which I’ll turn to in Section 2).

Russell’s sceptical hypothesis is an extreme version of a common philosophical approach. Such approaches begin by generating an epistemic demand: a philosophical bar is set, and knowledge-claims are checked to see if they can make the jump. Some claims, or sets of claims, might be high-jumpers, while others flop (and not in the Fosbury sense). We proceed by asking who makes it

[1] Russell’s thought experiment is intended only to demonstrate that induction cannot rest on logical, deductive grounds.

over the bar and who fails? Can we answer Russell's sceptic? These are grand, even gallant, starting places – but I think they aren't the only (nor the most productive) opening salvos we might make.

Perhaps we should start by asking whether there are fundamental differences between the past, present and future. If such fundamental differences are to be had, these could constrain or enable different kinds of knowledge or routes to knowledge. Here, we begin by considering the ontology of the past, in contrast to the present and future. Carol Cleland (2002, 2011) argues for something like this. She claims there is a physical asymmetry between the present's relationship to the past and the present's relationship to the future. Causes have multiple effects, and these spread through time. This multiplying aspect, according to Cleland, leads the present to *over*determine the past. That is, the way things are now is more than sufficient to guarantee the way things were, but not so the future, hence the asymmetry. Arthur Danto (1962) says something similar. The fixity of the past, for him, stands in stark contrast with the open future. And in virtue of this, our retrospective understanding of the past is simply of a different nature from our capacity (or lack thereof) to predict the future.

For Cleland, the temporal asymmetry underwrites the method historical scientists adopt. An approach that similarly ties together the relationship between the past and present, and scientific method, is *uniformitarianism*. The term was coined by the nineteenth-century geologist Charles Lyell (1837) and has undergone various – often increasingly complex – incarnations (see Gould, 1965; Camardi, 1999, although I'm basing my discussion on Rudwick's treatments, 1972, 2014). For our purposes, we can divide the idea into two claims: *actualism* and *gradualism*. Actualism claims that our knowledge of the past needs to be grounded in examinations of processes acting today: it is about how to get knowledge. Gradualism claims that processes in the past occurred at roughly the same rate as they occur today. Specifically, change is slow, incremental: great mountains and deep seas are formed by slow local changes; biodiversity is caused by patterns of individual survival, birth and death within particular populations. There are merely methodological or pragmatic versions of uniformitarianism that avoid substantive claims about the past, but such positions have nothing to say about the prerequisites of past knowledge.

Taking actualism and gradualism together, then, we have uniformitarianism. Actualism says that current processes are the key to past processes, and gradualism says that geological features are the result of those processes acting incrementally on grand scales; small changes scaled-up explain big changes: water trickles form mighty gorges.

Understood in a sufficiently weak way, actualism is important. To make inferences about the past, we need *traces*. We need to find remains – footprints, fossils, cave paintings, droppings and so on. Further, to make those inferences we need to understand the processes that shape the past, and examining current processes is one way of achieving this. Although I think taking traces as the basis of our knowledge about the past can be (and typically is) taken far too far (see Currie, 2018a, chapters 6-11), and that there are many ways in which past processes might differ from current processes, no doubt we have to start somewhere. Here, then, is a beginning for our examination of historical knowledge: 'Knowing about the past requires taking the present as having been shaped by its past—that is, to contain some kind of record which we can either decode, or perhaps decode one day'.[2]

Such a thought underwrites the trace-based reasoning we'll meet below.

Uniformitarianism's other half – gradualism – fares worse. Gradualism claims that change is slow and incremental; however, there are plenty of exceptions. Let's glance at a geological and then a biological example. 'Outburst floods' are enormous floods occurring when previously dammed lakes are freed. Such floods shaped North America's distinctive geology, often as a result of ice-age glacial blockades melting. This freed superlative amounts of water, which gouged out massive valleys and were partly responsible for the layers of soft and hard stone necessary for Badlands to form, as well as potentially changing weather patterns (Kehew & Teller, 1994). These are anything but gradual processes.

In biology, gradual speciation has been challenged in two ways. The most obviously gradual model of speciation is *anagenic*: one species gradually shades into another. We can contrast this with *cladogenic* speciation, when speciation occurs by 'splitting' two populations. There is considerable ongoing debate over to what extent speciation follows each process, but I'd be surprised if one model dominates (Plutynski, 2018). On a macroevolutionary scale, we can compare *phyletic* speciation – the thought that speciation occurs with gradual, steady change – with *punctuated equilibrium* (Gould & Eldredge, 1993). On the latter view, over evolutionary time most species are typically in stasis but when change comes, it comes dramatically. Again, there is debate about to what extent stasis or change is the dominant pattern in evolution, and to what extent speciation is gradual or punctuated. In light of these kinds of examples we can't just assume gradual speciation, nor, I think, can we assume gradualism more generally.

[2] Derek Turner would point out the explicit textual metaphor here!

Approaching the relationship between history and knowledge by starting with in-principle constraints on the possibility of knowing the past, or considering the relationship between the past and present abstractly, has a set of attendant risks. Such accounts often commit to a particular analysis of knowledge, or to particular metaphysical views, making their story about historical knowledge beholden to the fate of those commitments. I think we should avoid hitching our wagon to a particular analysis of the nature of knowledge or of the relationship between past and present. There is a danger of being fatally disconnected from the phenomena we're trying to understand. The hard-won, transformative historical knowledge that impressed us in the introduction was achieved via human ingenuity and sweat; it wasn't bequeathed from some general fundamental fact about the world, and the phenomenon holds independently of our solving philosophical puzzles about knowledge. The task of actually explaining how historical scientists successfully investigate the often obscure, often weirdly alien recesses of the deep past remains even if we've answered such questions. Under these conditions, striving for philosophical grounding can become a distraction, a red-herring, a roadblock to understanding.

I think a better approach starts with scientific practice: that is, we should examine what scientists actually do. Instead of doing philosophy first – setting a philosophical standard and examining practices to see if they meet it – we start with an examination of scientific work. How do historical scientists reason? What kinds of evidence do they provide? What knowledge-generating processes – fieldwork, experimentation, etc. . . . – do they engage in? After examining science, we see what philosophical systematization and lessons might be drawn from these practices. As such, we'll start by delving into some paleontology in the next Section.

You might complain: if my philosophical analysis begins with a descriptive case study, can I do the explanatory, normative work philosophy demands? If my criteria for good evidence are derived from descriptions of practice, don't I face a methodological dilemma? On one horn, my analysis is restricted to mere description, thus falling short of my normative goals; on the other horn, I fall afoul of the dictum that one mustn't derive an ought from an is. I think this complaint is mistaken: my discussion is neither purely descriptive nor normative and (or so I hope) avoids problematic circularity (Currie, 2015). To see why, I invite you to come along for the ride. In my concluding discussion I'll suggest that our reflections on the nature of the deep past and our knowledge of it are applicable to philosophy, and in a way that enables escape from this dilemma.

1.2 Growing Up Dinosaur

The archaeologist Christopher Hawkes saw historical texts as a crucial evidential crutch supporting old-world archaeology. Even if the target culture did not leave a written tradition, if it is in some way continuous with a culture which did, that continuity can underwrite rich reconstructions of cultural pasts: 'In rural economy, burial rites, technology, sociology or what not, there is always, somewhere or other, a point of reference within the historic order' (Hawkes, 1954, p. 160).

For Hawkes, insofar as we can rely on (and stretch) written records – find a point of reference – we can make inferences from material remains to specific, contingent features of past human societies. For the biology of the deep past, living descendants are the equivalent of historical texts. If an extinct critter has close relatives in the present, then examining those relatives can be an often powerful guide, providing a point of reference within living biota.

This is what makes dinosaurs so challenging.

Dinosaurs' closest living relatives are their progeny, the birds, and their cousins, the crocodilians. These lineages share a common ancestor around 240 million years ago, in the midst of the Triassic period (Green et al., 2014). Although birds and crocodiles provide some guidance for dinosaur palaeontologists, millions of years of evolution opened wide morphological, physiological and behavioural differences between these lineages. Neither living crocodiles nor birds, for instance, include in their ranks multi-ton, terrestrial, herd-living herbivores. If we wish to understand these critters, the horned and frilled ceratopsids, plumed Hadrosaurs, entanked ankylosaurs and earth-shaking sauropods, do we look to birds and crocodiles, or to their mammalian analogues (elephants and hippopotami), or to some combination of both? Are triceratops and friends more like scaled-up, wingless birds, or like reptiles playing at mammalhood? Are they somewhere in between? Are they something else entirely? To see these difficulties in play and how scientists respond, let's consider dinosaur 'ontogenetic development': their patterns of growth.

It is misleading to think that paleontology 'starts' with fieldwork, but our examination of dinosaur development has to begin somewhere. Fieldwork is not simply a matter of finding fossils. Specimen discovery is often challenging, and requires decisions about where to look and what is worth digging up. Fossil extraction is typically destructive. Fossils (particularly of larger animals) often weigh many tons, and don't typically hang out in convenient locations. Of those fossils that are retrieved, decisions must be made about storage and preparation. Preparing fossils is necessary for them to be analysed: only once the biological signal of the fossil has been split from the surrounding rock, can we discern and measure morphology (Wylie, 2015). And fossil preparation is an onerous task.

Post-preparation, decisions must then be made about which prepared fossils are worth analysing, how they should be stored and so forth. Each step, finding the fossil, deciding to dig it up, preparing it and then analysing it, require judgements about how to spend limited resources towards several – often conflicting – goals (Turner, 2016). Building data sets often involves specimens that were dug up and prepared with different practices and different questions in mind. And sometimes these differences make reconciling old data with new difficult (Wylie, 2017). Further, each step introduces a new possible source of bias in the eventual data (Wylie, 2019). The journey from discovery to evidential use is multi-stepped, and each step matters. Often a specimen having been extracted and prepared isn't sufficient for its use as evidence: palaeontologists adopt standards designed to preserve the authenticity and epistemic properties of fossil remains. Leading paleontological journals, for instance, do not accept new species on the basis of privately owned fossils, partly for ethical, partly for epistemic reasons (Havstad, 2019). Much of these complexities are encapsulated in Jack Horner and Mark Goodwin's investigation of how *Triceratops* grew.

Triceratops are instantly recognizable, three-horned ceratopsids from the North American Cretaceous. Although *Triceratops* specimens are relatively common (for dinosaurs), back in 2006 only four non-adults were described in published literature. This was partly because 'smaller *Triceratops* skulls and cranial elements were apparently overlooked, deemed highly incomplete or undesirable to collect' (Horner & Goodwin, 2006, p. 2757). Non-adult skulls are often more brittle, so less likely to survive the fossilization process, and those which do survive are typically incomplete. So if you're looking for the 'best' – most complete – skulls, then you'll focus on adults. The aims and standards of collecting affect what is collected (Wylie, 2017). Further, many of the *Triceratops* prepared in the past were prepared towards ends other than understanding their growth: ' … many previously collected *Triceratops* skulls in museum collections have undergone extensive restoration, are composites or lack contextual field documentation, making their use unreliable' (p. 2757).

In 1997, crews from the Museum of the Rockies and the University of California Berkeley began working the Hell Creek Formation in eastern Montana, aiming for a collection suitable for studying *Triceratops* development. The specimens were prepared under Horner and Goodwin's supervision, providing 10 full and 28 partial skulls.

Sorting the skulls by size and other signals of age, Horner and Goodwin hypothesized a four-stage sequence taking us from infant, juvenile, sub-adult to adult *Triceratops* (see Figure 1).

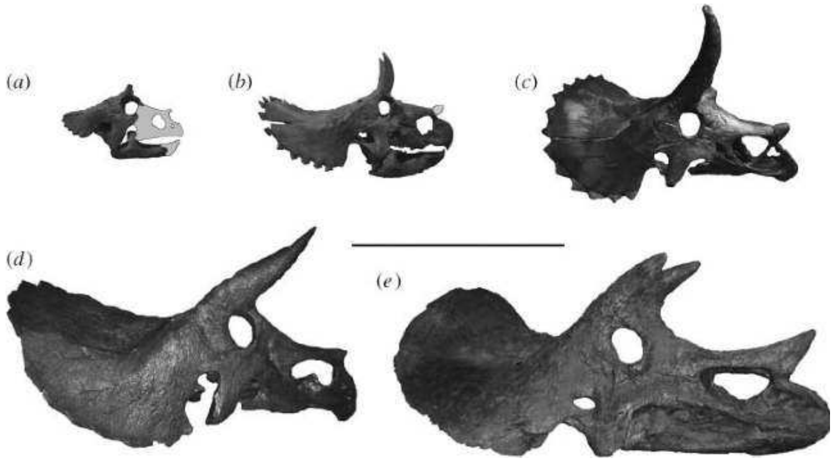

Figure 1 Horner and Goodwin's triceratops growth sequence. (Horner & Goodwin 2006, 2760) © Royal Society.

Even to uneducated eyes, the changes over a *Triceratops*' lifetime are striking. Most obvious are changes in the 'frill' (the bony crest jutting out the back of the skull) and in the 'post-orbital' horns (the two horns above the eyes). The frill begins rather unambitiously, jutting straight out the back in infants (Figure 1a), to increasingly splayed and dramatic across juveniles and sub-adults (Figure 1b, c). More strikingly, the post-orbital horns jut upwards with a slight backwards lean in early life stages (Figure 1a–c), but change dramatically in adulthood (Figure 1d, e), bending forwards over the eye.

Why does understanding *Triceratops* growth matter? What motivated nine years of collection, preparation, analysis and publishing? The reconstruction underwrites further speculation on Horner and Goodwin's part. They suggest that changes in post-orbital horns are 'probably visual cues of immaturity' (p. 2761). A quick glance at horn position would be a useful way for *Triceratops* to gauge age. More dramatically, in 2010, Horner, this time with John Scannella, combined this sequence with another *ceratopsid, Torosaurus*, to argue that the two were not different genera after all, but that *Torosaurus* represented the final adult stage of *Triceratops*. So, the proposed growth sequence provided a basis for further speculation and led to previous interpretations being re-evaluated.

Further, *Triceratops* skulls inform us about more than ontogeny. Scannella et al. (2014) used the same skulls to argue that *Triceratops* species evolved by gradual transformation. Triceratops in the Hell Creek formation have two morphotypes: *T. horridus* and *T. prorsus*. Instead of arranging skulls by life-stage, Scannella and company arranged them by stratigraphic sequence (that is, organized by the layer of rock in which they were discovered). Because strata