Studies in History and Philosophy of Science 62 (2017) 14–21

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

journal homepage: www.elsevier.com/locate/shpsa

In defence of story-telling

Adrian Currie [a,*], Kim Sterelny [b]

[a] Centre for the Study of Existential Risk, David Attenborough Building, Pembroke St, Cambridge CB2 3QY, UK
[b] School of Philosophy, RSSS, Coombs Bld, Fellows Rd, Australia National University, Acton, Canberra, ACT, 0200, Australia

ARTICLE INFO

Article history:
Available online 17 March 2017

Keywords:
Historical reconstruction
Narrative
History
Storytelling

ABSTRACT

We argue that narratives are central to the success of historical reconstruction. Narrative explanation involves tracing causal trajectories across time. The construction of narrative, then, often involves postulating relatively speculative causal connections between comparatively well-established events. But speculation is not always idle or harmful: it also aids in overcoming local underdetermination by forming scaffolds from which new evidence becomes relevant. Moreover, as our understanding of the past's causal milieus become richer, the constraints on narrative plausibility become increasingly strict: a narrative's admissibility does not turn on mere logical consistency with background data. Finally, narrative explanation and explanation generated by simple, formal models complement one another. Where models often achieve isolation and precision at the cost of simplification and abstraction, narratives can track complex changes in a trajectory over time at the cost of simplicity and precision. In combination both allow us to understand and explain highly complex historical sequences.
© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

In the early 19th Century, a cache of 78 ancient chessmen, mostly carved from walrus ivory, were discovered on the Scottish island of Lewis. They are pictured below (Fig. 1), with a provocative quote.

> Who carved them? Where? How did they arrive in the sandbank—or, as another account says, that underground cist—on the Isle of Lewis in western-most Scotland? No one knows for sure: History, too, has many pieces missing. To play the game, we fill the empty squares with pieces of our own imagination. (Brown, 2015, 1–2).

This quote demands a narrative: an explanation which follows the causal trajectory of the chessmen's origin and subsequent history. Such narratives are common in both historical and scientific reconstruction of the past.[1] Nancy Marie Brown's recent popular history *Ivory Vikings* combines two narratives about the Lewis Chessmen. The first story covers the last few centuries, detailing debates between art historians, archaeologists and antiquarians about the provenance, manufacture, and purpose of the pieces. The second story is set in the 9th to 13th centuries, and focuses on the social, cultural and economic world of the Lewis Chessmen: the medieval North Atlantic. Brown's emphasis on the role of imagination—story telling—is apt for both narratives. In uncovering history, we draw on material remains, such as the those of the economic and social lives of these communities, and the chessmen themselves, as well as surviving literature like Iceland's rich sagas and hints in the linguistic patterns of contemporary Scandinavian languages—a tapestry of evidence. This evidence is fuel for narrative explanation; stories of how and why the pieces were made, and how they ended up where they did. In developing narratives, imagination plays an important role, as the passage of time erodes elements in the chain of causation; there are 'empty squares' that our imagination must fill.

It is our contention that such story-telling is central to successful historical reconstruction, and moreover that there is no reason for blanket scepticism about such reconstructions. Further, we argue this is just as true for science as it is for history. In this regard, practitioners of human history are methodologically continuous with archaeologists, geologists, cosmologists and palaeontologists. There are differences of course: historical scientists tend to be more concerned with understanding general patterns than historians. They seek to identify general mechanisms that shape causal trajectories through time; for example, the features that determine extinction risk in periods of mass extinction, and those that

* Corresponding author.
E-mail addresses: Adrian.currie@ucalgary.ca (A. Currie), Kim.strerelny@anu.edu.au (K. Sterelny).

[1] For general discussion of historical science, see Cleland (2002), Currie and Turner (2016).

http://dx.doi.org/10.1016/j.shpsa.2017.03.003
0039-3681/© 2017 Elsevier Ltd. All rights reserved.

Fig. 1. Some of the British Museum's Lewis Chessmen. Including King and Queen (front centre), knights (back row), bishop (centre), pawns (front ends) and rooks (middle ends). The rook on the right is biting his shield - traditional Berserker behaviour. (source: Wikimedia commons).

determine extinction risk in less dramatic times. But, like historians, they also aim to explain individual historical episodes of particular interest or importance: the formation and breakup of Pangaea; the radiation of flowering plants; the development and spread of agriculture.

Part of the explanatory agenda of historical science involves the identification of similarities between historical trajectories: as noted above, both the general and the particular is of interest to them. The biotic recolonization of Krakatoa after the eruption of 1883 might be similar to the re-establishment of ecosystems on other volcanic islands after eruptions in informative ways (Thornton, 1996). But there are non-trivial differences as well, and ecologists are interested in both the differences and the similarities. In contrast, while the oxidation of iron is a process that takes place in space and time, and at varying rates depending on local conditions, chemists are not typically interested in, say, the specific series of events occurring as an abandoned car shell rusts into the soil. For us, a narrative is a specification of an individual trajectory of this kind.

As we shall understand them, then, a narrative is a candidate explanation of a particular causal trajectory in the past thought to be of interest in its own right. Narratives are not mere chronicles—they do more than provide an ordering of events. They posit links—often causal—between them;[2] earlier events conspire to produce later events. This account of narrative leaves much open. Obviously, narratives can be more or less detailed. Likewise, narratives may present events as being more-or-less contingent.[3] In principle, a narrative explanation of the origins of World War I might be given precisely to underscore its inevitability, charting a perfect storm hitting the European political system early in the twentieth century. We also leave open the possibility that a narrative explanation of a particular historical episode might be intended to illustrate some general mechanism or tendency. A narrative explanation of the origins of World War I might also be intended to illustrate the threat to peace posed by political systems involving great powers and competing alliances. However, the narrative must intend to capture and explain, at some level of grain, the specific features of that trajectory; the features that make it of genuine interest.[4]

We take it as obvious that historians and historical scientists construct narratives. Our aim is to defend the epistemic viability and productivity of this practice. Building a narrative might seem unproductively speculative, because a narrative typically involves the reconstruction of causal intermediaries that have left no unambiguous trace in the present; positing rather than finding links in a causal chain. Since narrative explanations explain via these causal chains, the explanation as a whole is persuasive only if the identification of each link and its causal connections is persuasive. That is why the charge of "story-telling" is potentially serious; leading to a supposed contrast between the 'real' science, the more-or-less firmly established links between material remains and the past, and 'mere' storytelling, the construction of imagined links between those pockets of evidential confidence.[5]

Narratives can be problematic in two ways. First, because they are intended to specify what is distinctive about a *specific* trajectory, we cannot take one instance as a model of them all (see Tucker, 1998). Second, the dispersal and erosion of evidence about the sequence—the information-destroying processes of decay—often

[2] We take narratives to be causal (though not all do), but we will not argue that here. Likewise, we will remain neutral on the nature of cause and causal explanation.

[3] *Pace* John Beatty's view (2016).

[4] Our account of narrative is intended to be more-or-less consistent with others in the literature. One of us Currie (2014) has previously endorsed an extremely thin notion of narrative, identifying it with the explanation of token events; the other Sterelny (2016) has identified narrative with explanations with particular modal properties. Other philosophers (Beatty, 2016; Hull, 1975; Roth, 2008) provide accounts of narrative which do not depart from our account in ways which matter to our argument.

[5] The 'story-telling isn't science' stance is most often expressed in casual conversation, but it is expressed in Aunger's (1995) discussion of skepticism about ethnographic reports, Herrick's (2004) position that science is not 'objective' but rather provides 'narrative coherence', the apparent conflict between 'narrative' and 'evidence-based' approaches to medicine (Meisel & Karlawish, 2011 discuss, but do not endorse the conflict) and those biological scientists decrying 'just-so stories' (starting from Gould & Lewontin, 1979). For a quite different defence of the role of storytelling in science, see Grobstein (2005).

leave hypotheses locally underdetermined[6] (Turner, 2007; Sober, 1988). This dispersal is not typically uniform: we often have relatively reliable and direct evidence about some episodes in the sequence, and much less reliable or direct evidence about others (and about the causal connections between them). Building narratives relies, we shall suggest, on some form of story-telling to forge the links between these more initially evidentially secure elements of the overall trajectory. Narrative, insofar as it involves such reconstructions, it is thought, may play a heuristic role in guiding the historian, but narratives are speculative in problematic ways, and the reconstructed elements of narratives are insufficiently constrained by the more secure parts. It is true (the suspicion runs) that these reconstructed elements must cohere with the trace-based elements. But coherence (a merely logical notion) is a weak constraint. Too many potential narratives could equally suit our evidence.[7]

Carol Cleland notes:

> If a historical hypothesis … draws its explanatory power primarily from the coherence and continuity of a quasi-fictional story, then historical natural science really does seem inferior to experimental science; in the absence of empirical warrant a narrative explanation amounts to little more than a ‘just-so’ story (Cleland, 2011, 17).

Of course some narratives are poor history; poor historical science. They really are appropriately mocked as “just-so” stories. But we think a general suspicion of narrative-based explanation is misplaced. We shall argue in section two (i) that reconstructed narratives need not be problematically speculative, and in historical reconstruction they typically are not problematically speculative. Further, (ii) speculation often benefits historical reconstruction by guiding the identification of further evidence. In section three we argue (iii) that coherence typically imposes quite tight constraints on narrative explanations. Moreover, (iv) there is no fundamental epistemic distinction between our identification of episodes in a chain of causation for which we have direct, trace-based evidence (so for example a layer of shocked quartz and iridium signals a large bolide impact at or near the K/Pg boundary) and those episodes we reconstruct from indirect evidence (extinctions shortly after the K/Pg events). Finally, in section 4 we shall suggest (v) that narrative explanations themselves impose constraints on more general models of historical processes. There is a fruitful interplay between general, formal approaches to the past and the more particular and discursive approach exemplified by narrative. In sum, storytelling plays an important and central role in successful historical reconstruction.

2. The supposed vice of speculation

One worry about narrative is that it involves speculation—that it outruns our available evidence in problematic ways. We will distinguish between idle and productive speculation, and suggest that the latter solves an important problem in reconstruction: identifying relevant evidence. Further, productive speculation serves to increase the empirical constraints on historical reconstruction, as we shall see in our discussion of coherence in the next section.

The fact that a hypothesis is speculative is—at least often—no objection to its scientific legitimacy. Speculation can be a vice if it is disguised, idle, or a Trojan horse for unjustified preconceptions. Indeed, the idea that speculation is such a Trojan horse seems to explain Henry Gee’s rejection of narrative in palaeobiology. As he sees it, the sparse and patchy data of history make narrative explanations likely to reflect our preconceptions back at us.

> … the scale of geological time that scientists are dealing with … is so vast that it defies narrative. Fossils, such as the fossils of creatures we hail as our ancestors, constitute primary evidence for the history of life, but each fossil is an infinitesimal dot, lost in a fathomless sea of time, whose relationship with other fossils and organisms living in the present day is obscure. Any story we tell against the compass of geological time that links these fossils in sequences of cause and effect—or ancestry and descent—is, therefore, only ours to make. We *invent* these stories, after the fact, to justify the history of life according to our own prejudices … Geological time admits no narrative in which causes can be linked with effects (Gee, 2000, 2).[8]

Gee’s suggestion seems to be that narratives are not just underdetermined by the data; but that under-determination is deep, unimprovable, and often unnoticed or under-estimated. We think that Gee has probably identified a genuine danger. The human mind likes stories (Boyd, 2009) making the step between finding a story attractive, and taking it as true, all too easy. Stories can be seductive, so it is important for historians and historical scientists to be methodologically self-aware; to be continually concerned with testing their narrative explanations. But we do not think that methodological self-awareness is typically absent in the historical sciences. Indeed, Alison Wylie has argued that at least in archaeology:

> … archaeological evidence has a striking capacity to disrupt settled assumptions, redirecting inquiry and expanding interpretive horizons in directions no one could have anticipated (Wylie, 2011, 371).

While it’s always possible to be captured by one’s own stories, we see no reason to accept that local under-determination is intractable, even in Gee’s special case of ancestor-descendant relations in the fossil record. (See, for example, recent doubts that the australopithecines were hominin ancestors: White, Lovejoy, Asfaw, Carlson, & Suwa, 2015).

So speculation is a vice—is idle—when it is pointless: when it cannot or does not productively direct further inquiry; when it is not used to construct alternative scenarios to guide a search for evidence which would favour one at the expense of the other. There are such speculations in the historical sciences; for example, attempts to interpret the specific meaning (as distinct from the functional significance) of the material symbolism of long-extinct cultures. See for example Martin Porr’s work (2015). Porr’s speculations about Palaeolithic figurines are in fact quite plausible, but it is hard to see how he could use them as a guide to further test his interpretations. Another example is the thought that sauropod

[6] Hypotheses are *locally* underdetermined when the currently available evidence is insufficient to distinguish between them (as opposed to the hypotheses having identical empirical consequences).

[7] These complaints are not often found in the published literature, but both of us have met it regularly in conversation, and one of us regularly in referee’s reports on his narrative-based explanations of hominin evolutionary history.

[8] It’s worth pointing out that Gee’s complaints are specifically about identifying fossil taxa with living taxa, claiming, for instance, that some extinct lineage is the ancestor of some extant lineage (as opposed to a sister-taxa). And indeed such claims may be problematically speculative—but are not so in virtue of their narrative quality.

dinosaurs had a unique thermoregulative system which switched between endothermia and ectothermia in ontogeny (Farlow, 1990).[9] This suggestion solves some otherwise confusing aspects of sauropod life-ways (see Currie, 2016), but has not as yet led to further studies. Such speculations have not thus far been coupled with suggestions as to how they can be tested and refined, and so are so far empirically idle,[10] and perhaps open to the worry Gee highlights.[11]

However, we think unproductive speculation is mercifully rare in science—and indeed often knowledge generation is organized to avoid it. For instance, palaeontologists use different systems of categorization for trackways and for bodily fossils. Fossilized bones and trackways typically record different information at different grains.[12] Most of the time, attempting to identify what critter left which tracks is idle—not only do we lack a method of supporting such hypotheses, but they fail to generate new lines of evidence. When Xing et al. (2013) identified swimming theropod trackways (the tips of claws dipping into the ancient riverbed) they identified them with the ichnogenus *Characichnos,* not with any particular theropod taxa—neither *microraptor,* nor *T. rex,* nor any other. The use of parataxonomies[13] insulates palaeontologists from the problematic—idly speculative—idea that one might connect a particular trackway to an extinct lineage at as fine a grain as the species-level.[14]

Productive speculation, by contrast, solves a pervasive problem in historical reconstruction: identifying evidential relevance. Overcoming under-determination in historical reconstruction requires a wide variety of evidence sources, and it is often difficult to identify these sources prior to investigation. Productive speculation provides the scaffolding necessary for progress in the face of history's opacity. Xing et al. suggest that the trackways indicate a fairly regular swimming pattern in theropods—hypotheses potentially amenable to further biomechanical probing. Further, such single cases can themselves be compared and contrasted to underwrite further hypotheses (and rather systematically, see for example Lockley, Xing, Kim, & Matsukawa, 2014). Moreover, while there is probably little point in attempting to unify paratoxonomies based on trackways and those derived from bodily remains, this doesn't mean that the two lines of evidence cannot be fruitfully integrated. Working out how extinct lineages walked, for instance, often relies on tying together both theories of the physiology and anatomy of gait and anatomical reconstructions on the basis of fossils and trackways—and the resulting hypotheses are often probed using simulations (Turner, 2009). On the basis of Xing et al.'s (2013) hypothesis that *Characichnos* tracks represented a swimming therapod, they were able to further model how such creatures swam. Sellers, Margetts, Coria, and Manning (2013) constructed a simulant sauropod on the basis of anatomical and muscular speculation. Examining the simulation, they generated a prediction about how sauropods walked. This prediction was borne out via examination of sauropod trackways. It was in virtue of the hypotheses which the simulation generated that the trackways were evidentially relevant. Speculation, then, reveals avenues for testing and scaffolds further investigation (Currie, 2015).

This point applies in human history as well. Brown's narrative highlights and connects two well-confirmed hypotheses about the past. That a large set of chess pieces were constructed from walrus ivory somewhere in the Norwegian sphere of cultural influence in the medieval North Atlantic; and that a skilled Icelandic craftswoman (mentioned in the saga of Bishop Pall) named Margret the Adroit was active in the early 13th Century, and is identified with a fine ivory crozier discovered in Pall's tomb. Brown attempts to convince us that the two are connected—that Margret herself carved the figures. This is storytelling in our sense: Brown posits a causal relationship between two events, via an intermediary link for which we have no direct evidence, thus shifting from a mere chronicle (that is, a temporal ordering) to a history proper.

Historical reconstruction often proceeds by identifying and overcoming local under-determination —and overcoming under-determination requires locating evidence. On the discovery of the Lewis chessmen, it would have been very difficult to predict the features of the North Atlantic that would matter in reconstructing their history. It is only, for instance, in light of the hypothesis that they were made in Iceland—a hypothesis which itself depends on evidence concerning the ivory trade from Greenland and the craft and material capacities of medieval Icelandic culture—that our knowledge of Bishop Pall and Margret the Adroit is evidentially relevant. As noted just above, one way speculation can be productive is when it guides the search for relevant new facts. But facts are relevant as and when they impose tighter constraints of coherence on suggested narratives. So our response to scepticism about speculation depends on our view of the importance of coherence.

Is Brown's narrative incurably speculative; and if it is, is this empirical weakness representative of narrative reconstruction which fills in links between secure, established episodes? The sceptic suspects that while the identification of Margret as the carver coheres with the known facts, such coherence is too weak a constraint. There will be many equally good narratives, and no productive way of showing that one is more probable than the others. We disagree. Coherence is not mere logical consistency with a few known facts, and so plausible narratives are not so easy to come by. In the next section we develop these claims.

3. In defence of coherence

Coherence is a much-under-rated epistemic virtue. Achieving it involves much more than establishing mere logical consistency between what is said about one stage of a trajectory and what is said about the other stages. If a narrative of the re-colonisation of Krakatoa or the making of the Lewis chessmen is coherent, it has satisfied many empirical and theoretical constraints. In the case of the Lewis chessmen, our narrative must of course avoid human impossibilities (virgin births; hale and hearty 150-year olds) and improbable co-incidences (identical twins separated at birth meeting on a desert island). But more seriously, the agents and their interactions have to be of the kind made available by the social, technical, ideological and economic resources of the medieval North Atlantic world. Brown's identification of Margret the Adroit depends, for example, on high-end walrus-ivory carving being a rare skill; there were few in that world who could have made those pieces. Constraints like this are often difficult to simultaneously

[9] Presumably this is not idle speculation in principle: it strikes us that hibernation and related behaviours could provide inroads to the mechanisms behind changes in thermoregulation.

[10] Such speculation might be justified in contexts when how-possibly explanations are called for, as in some adaptationist explanations of complex traits.

[11] Gould & Lewontin's complaints about adaptationist reasoning is in part clarified by this distinction: the charge of 'just-so' storytelling is in effect the charge of idle speculation: adaptationist hypotheses fail to open new investigative routes and actively discourage them (here is not the place to consider whether such a charge is plausible).

[12] See Turner (2007) and Finkelman (2016) for philosophical discussion of paratoxonomies.

[13] Here 'parataxonomy' refers to taxonomies tracking different kinds of trace (i.e., fossilized bones versus fossilized trackways). We gather (thanks to a referee) that this term is sometimes used to indicate alternative taxonomies constructed by, say, amateurs and professionals.

[14] Palaeontologists call non-body fossils such as trackways or burrows 'trace-fossils', we are using 'trace' in a more general way here, as referring to a present outcome of a past event or process.

satisfy; despite their best efforts, deliberate historical fictions often fail it, as historical novels often project contemporary attitudes and responses into (for example) early nineteenth century agents.[15] Analogous constraints are relevant to narrative reconstruction in the historical sciences: a reconstruction of the greening of Krakatoa has to be consistent with obvious general principles of ecology (no herbivores before herbs) but also with many specific factors that characterise the local region. The supposed early pioneers must be available within the regional biota; the dispersal mechanisms must be independently credible; the conditions on Krakatoa after the eruption (as attested by geochemistry) must be within their known physiological tolerances, and so on. In evolutionary biology, phylogeny provides an increasingly powerful constraint on candidate narratives, as information about the timing and branching patterns of major clades becomes increasingly available.

In the rest of this section we show how tightly these constraints, in favourable cases, constrain potential narratives, and how our understanding of these regional (and sometimes global) mechanisms lessen the epistemic significance of the distinction between episodes which leave a direct trace, and episodes whose character must be reconstructed indirectly.

The key idea is that the significance of a trace is itself inferential. When a narrative specifies a sequence of events in the past—a pyroclastic eruption; a smoking mound of volcanic debris; a mound cooled and moistened by rain; a mound with an initial dusting of organic materials; the arrival of small spiders and other small insects on the wind; a little erosion and soil formation—we should not divide these events into those which are directly attested by their surviving signature in the present, and those which are merely inferred or imagined. For all traces need to be interpreted in the light of often complex and sometimes controversial middle range theory; theory that tells us how an event's footprint at a time is made and then transformed.[16] There are of course massive differences in reliability: we can very reliably infer from megalodon fossil teeth to the presence of a very large predatory shark in the oceans of the recent past. The tooth is a trace, and it is a very reliable signature of a large predator. We also have reliable evidence that the megalodon is no longer with us, and has not been since the Pliocene or early Pleistocene. That knowledge about the past is not based on a trace; rather, it is based on our failure to find traces in many deposits where we might expect them, had the megalodon existed at later times. While some inferences are very reliable, others are less so: inferring the megalodon's hunting strategy is less secure, even though those teeth provide clues, especially when combined with robust patterns in shark behaviour. Even so, our knowledge of the shark's existence is more secure than our knowledge of its behaviour. That said, inferring events in the past on the basis of their material remains is not different in kind, epistemically speaking, from less direct inferential strategies. We have knowledge of hominin diets from direct traces: from the nature of teeth and jaws; from isotope studies of bones. We have knowledge from indirect sources: from wear patterns on stone tools; from middens and other remains; from inferences from skeletal remains to estimates of the shape and volume of gut tissue; from calculations about energetic demands imposed by hominin morphology and developmental biology. These evidential streams do not differ in a principled way *vis-à-vis* their reliability or their dependence on middle range theory. So while the inference from a trace to its historical cause is sometimes very reliable indeed, traces do not give us theoretically unmediated, observation-like access to the past. Thus, while some episodes in a trajectory are identifiable with greater reliability than others, we shall suggest that the differences are mostly differences of degree.

We insist that coherence often imposes tight constraints on potential narrative explanations; under those circumstances, the production of a narrative is itself an epistemic achievement. We will begin our defence of that claim with a reminder of how rich, and richly enmeshed in our knowledge of general mechanisms, our knowledge of the past can be.

Consider, for example, the discovery of a fossil bone: say, a tibia. From this discovery, we can infer the existence of the bone's owner, for we understand how, over long periods of time and under specific conditions, bones can be re-mineralized while retaining their structure. Moreover, we can identify the owner of the bone as a vertebrate, as that anatomical structure is only known in vertebrates. And from this, we can infer that the owner of the tibia also had a fibula, again through our knowledge of vertebrate anatomy. The structure of the bone and histology might yield further clues about the bone's owner: the fine structures of mammal bones differ from those of birds, reptiles and other non-mammalian animals. Often anatomical structure allows relatively safe inferences about whether the animal was bipedal or quadrupedal. From information about anatomy we can often infer gait, due to stable regularities between these features (Davis, 1964). For instance, an equal ratio between forelimb and hindlimb length signals a cursorial gait in quadrupeds. Moreover, features of anatomy, ancestry and gait are suggestive of features of the animal's physiology—its thermoregulation, its energetic demands and the like. Well-resolved phylogenetic analysis places the organism into ancestral context, as well as underwriting further reconstruction via the comparative method. Teeth, in particular, often carry phylogenetic as well as functional information. The fossil's stratigraphic placement can carry information about its age and about its palaeoecology, though again only with the assistance of a rich set of middle range theories about site formation and taphonomy, and the physics of radioactive decay and other dating methods.

Thus fossils often carry informationally rich signals from the past, but only do so in conjunction with a complex set of bodies of knowledge, ranging from information about local environmental features, to regularities across phylogenetic groups, to physical and chemical processes of decay. Indeed, the fact that a fossil is the mineralized bone of a once living animal is itself an important and unobvious discovery (Rudwick, 1972). In making these points about palaeobiology and palaeoecology, we are not (yet) building a narrative. Rather, our point is that the richer our general picture of a causal milieu, the more constrained any narrative hypothesis about a trajectory through that milieu will be. The library of plausible stories shrinks rapidly. The historical sciences of geology and palaeobiology have developed, and are further developing, increasingly rich, and temporally well-resolved models of the past's causal structure, especially those of the relatively recent past. We now have quite detailed models of the palaeoecology and climate of the Pliocene and Pleistocene world, and these sharply constrain the narratives we can give of (say) hominin evolution. For example, there is unmistakable evidence of major dietary changes beginning a little more than two million years ago: with reductions in tooth and jaw size and reductions in the musculature powering those jaws, while at approximately the same time relative brain size (and brains are energetically expensive tissue) began to increase. The nature of this change remains controversial (though probably involving some mix of increases to both meat consumption and food processing) but the signal of a major change is not controversial. Any narrative of hominin evolution needs to be integrated

[15] For example: Matthew Hervey, the hero of Allan Mallinson's enjoyable series based on the British Army of the Napoleonic War is suspiciously free of the class and ethnic prejudices of the time.

[16] The term 'middle-range' theory is from the archaeologist Lewis Binford (1977) and adapted by Peter Kosso (2001).

with these and many other aspects of hominin morphology, physiology, behaviour, and distribution in space and time. The rich and varied evidential streams historical scientists exploit constrain the space of plausibility. Generating good narratives under such circumstances is a significant epistemic achievement.

Time to take stock. Here is a simple, generalized account of an historical reconstruction's warrant. We understand the processes which shape history. Fossilization, political revolutions, mineralization, mass-extinctions, economic pressures, and so forth, have more-or-less regular effects. Moreover, the signs of those effects change over time in reasonably recognisable and well-understood ways. Reconstruction of the past is possible in virtue of these processes and our understanding of them. Typically, when people consider the evidence underlying reconstruction, they think of it as a causal web linking some past event to the present. Processes of mineralization and taphonomy link a fossilized bone, a current trace, to an extinct animal; and our understanding of fossil formation, site transformation, and biology gives us some understanding of that animal. But these causal webs also link events in the past to one another (Currie, 2016). In one of the historical sciences' most famous cases, there is a causal web linking an impact crater, an iridium layer, shocked quartz, and tektites to regional and global environmental change, and to the extinction of a spectacular group of animals. The role of these events in the extinction remains controversial, but any narrative of biotic turnover at the K/Pg boundary must incorporate these elements.

What is true on this grand scale is likewise true on a smaller scale. Brown connects the Lewis chessmen to Margret the Adroit: the latter carved the former. The basis for this involves an understanding of many local facts about Iceland's history—these facts play the same kind of role as fossilization, stratigraphy and comparative anatomy in reconstructing extinct animals. Brown identifies Bishop Pall as a likely sponsor for the chess pieces. His emphasis on beautifying the church with music and architecture (rather than delivering radical sermons) is specified in the saga bearing his name, and reinforced by archaeological finds. Moreover, his ideological stance towards church and state matches that implied by the chess-set. According to Brown, more radical bishops would be unlikely to invest in a version of chess which relegates religion's power to such a subsidiary role. Iceland's importance in the trade of walrus-ivory from Greenland is well documented both in sagas and archaeology. Specific correlations between historical Icelandic individuals (such as Queen Gunnhild the Grim) and the carved figures further link the finds with Iceland.

Brown's reconstruction is far from certain. But just as middle-range theories of taphonomy support the inference from fossil tooth to extinct animal to its palaeoecology, the confluence of evidence generated by our knowledge of the Medieval North Atlantic connect Margret the Adroit to the Lewis Chessmen. And, potentially, this connection could itself form the basis of (or at least inform) further conjectures about the past. For example, one might attempt isotope or other chemical analyses of the chessmen, or analysis of the artistic styles and techniques, to link that ivory to other potential examples of Margret's work. A well-established link to Margret, and thus Iceland and Pall, could lead to further ideas about the purpose of the chess-set (were they, for instance, a gift involved in solidifying political alliances?).

The construction of a narrative, then, is a considerable epistemic achievement: our rich knowledge of the past's causal structures provides strict constraints on admissibility. We think that this conclusion about narrative explanation is true of explanation more generally. As our information about the causal background is enriched, coherence becomes an increasingly important, increasingly demanding constraint. So, for example, a theory of the stability conditions of human cooperation has to fit a large number of empirical and theoretical constraints. When highly constrained narratives are also productive, when they underlie further testing and identify further evidence, they also extend our reach into the past.

4. Narratives and models

In the last two sections we have, in effect, fended off claims that narrative explanations fail to live up to the epistemic standards of good science. Indeed, story-telling extends our reach into the past, overcoming under-determination by enabling the identification of relevant evidence. In this section, we make explicit an idea that has so far been implicit: narrative explanations add something methodologically central to the historical sciences, specifically, they complement formal models and quantitative techniques.

We should clarify what we take this complementary relationship to be. In the context of historical reconstruction—the process of uncovering the past—narratives and models can be understood as performing different roles which match their distinct theoretical virtues. Formal models in the historical sciences enable theorists to explicitly represent and assess varying hypotheses about the relative importance of different factors, and to represent different scenarios in which the same factors operate, but with different relative strengths and across different background conditions. Paleoclimatological models probe the sensitives between global temperatures and atmosphere in the deep past; biomechanical simulations can generate and support hypotheses about the gait of extinct critters. But formal models are less good at capturing complexity and in representing how background factors change as history unfolds. Models face trade-offs between complexity and tractability, and are forced to make simplifying assumptions.[17] Narrative explanations can incorporate complexity: as we saw above in discussing coherence, as our picture of the environment in which history is unfolding becomes richer and more detailed, the narrative explanations of those dynamics become more tightly constrained. A bug for one approach is a feature for the other. Just as models make explicit relative importance amongst a narrow and tightly specified set of factors, typically acting in somewhat idealized conditions. Similarly, narratives can and should make explicit the complexity of causal trajectories and the interaction between a change in a focal phenomenon (for instance, the evolutionary emergence of distinctive forms of human cooperation) and the environmental and evolutionary background in which that trajectory takes place. They do so though, at the cost of quantitative precision. This is one respect in which the methods are complementary: one picks up the detail and specificity that the other almost inevitably sacrifices in search of generality. In other contexts, models and narratives may be further integrated,[18] but we will focus on their complementarity.

We just mentioned human cooperation for a reason. The question of how our lineage evolved our distinctive, complex social worlds is one which lends itself to both narrative and model-based probing. In effect, a narrative explanation of the emergence of human cooperation provides a causal sequence linking base-line conditions—great-ape-like social worlds—to extensively cooperative arrangements such as our own, wherein cooperative and collective action is obligatory. Recently, Boehm, Tomasello, and Sterelny have all published narrative-style explanations of the emergence of such cooperation (Boehm, 2012; Sterelny, 2012; Tomasello, 2014). These narratives all depend on formal models

[17] For general discussions of the trade-offs in modeling see Michael Weisberg (2013).

[18] See, for instance, Currie (2014) discussion of 'simple' narrative explanations.

to isolate and explore the potential causal interactions between different features of these changing worlds. But nonetheless, all three candidate explanations are narratives. By contrast, Sam Bowles and Herb Gintis's The Cooperative Species (2011) is a sophisticated and impressive attempt to give an account of the evolution of human cooperation, and is no narrative—rather, a series of formal quantitative models are used. There is no attempt to trace a trajectory of changes from our ancestors to ourselves. And nor is there a "master model" in which the various factors which combined to make human cooperation possible — social learning and teaching; cooperation with respect to inter-group competition; the evolution of norms and punishment; reciprocation; reproductive cooperation — are all represented. Such a "master model" would be intractable. We will discuss their project in some detail, not because we think poorly of it, but because it is such an impressive instance of its kind; its clarity and power serve to show the limits of trying to do without narrative altogether, in giving an explanation of a complex, multi-stage transition.

Instead of a narrative, Bowles and Gintis try to show that the evolution of each element in the cooperation stew is plausible, factor by factor. In effect, the various ingredients of human cooperation are treated formally in isolation. The strategy is to take a series of distinctive features of human social life — helping, punishment, responsiveness to norms, intergroup aggression — and develop a cluster of models that represent the emergence of each feature, typically holding fixed a set of background conditions, but equally typically showing that the trait's emergence does not depend on very particular, or implausibly extreme, parameter values. Their nuanced, complex picture rules out the option of a master model in which all these ingredients are represented, and a set of plausible trajectories generated. Instead, the emergence and stability of each ingredient is modelled separately and repeatedly.

This model-based strategy, eschewing a narrative, and without a master model, means that there is no explicit representation of the coordinated evolution of the suite of social, cognitive, communicative, demographic and motivational factors that ultimately made our extraordinary levels and kinds of cooperation possible. That is a gap in itself: there is no integrated step by step account of the transition to human ultra-cooperation. In principle, a set of models of this form could be used, in conjunction with historical data, to construct a constrained narrative, with the narrative account of change in each key aspect of human social life tightly constrained by the appropriate model in the set. In practice, this conception understates the complementarity of the two approaches. A narrative can make salient hidden constraints on the models. To see this, consider Bowles & Gintis' approach to punishment.

The standard view of punishment is that while it is easy to explain the stability of collective punishment it is difficult to explain its origins. When the willingness to punish is common, it is easily maintained as cost is low: because punishment is rarely necessary, and the costs are divided amongst the participants. Conversely, when it's rare (as all traits initially are), it is very expensive, for freeriders have not yet learned to desist through fear of punishment, and because the cost is spread across the few rather than the many. Bowles and Gintis dissent from this line of thought. They suppose that an agent's willingness to punish is sensitive to her assessment of its cost and effectiveness. So punishment invades as a conditional, threshold-dependent strategy. Punishers punish only when (perhaps initially through chance) they reach a threshold in the local environment. This strategy allows their frequency to grow, since coordinated punishment is not too expensive, and the threat of punishment, allied to the knowledge of past punishment, induces non-punishers to cooperate. So punishers are compensated for the costs of punishment by the rewards of cooperation, and since punishers enforce cooperation through punishment only when those willing to punish are locally common, most of those rewards of cooperation go to those who signal their willingness to punish.

The idea, then, is that punishment invades as a conditional strategy, as a signal and response system that enables players using it to benefit from cooperation. A problem for this analysis arises when it is embedded in a larger narrative. If the control of freeriding is critical in establishing a cooperative social environment, and if freeriding must be controlled by punishment, coordinated punishment must evolve early in the transition from great ape to sapiens-like social worlds.[19] But these models also assume capacities to signal, to interpret signals and to coordinate in inflicting punishment; capacities that would only evolve late, because they evolve only in a social environment that is already much more cooperative than those revealed by great ape ethnography. Conditional punishment is too sophisticated to be an early and foundational form of social behaviour. Tracking a punishment threshold depends on active and reliable signalling and interpretation, as conditional punishers census their local density. We expect signal-comprehension-coordination capacities to evolve incrementally in a cooperative world; they cannot be assumed for free as an explanation of the origins of such a world. Moreover, the models assume that non-punishers cooperate in response to punishment, rather than counter-punish. This assumption is far from trivial, for in experimental games, punishment quite often attracts counter-punishment rather than cooperation (Herrmann, Thöni, & Gächter, 2008; Gächter & Herrmann, 2009). The best guess is that punishment only induces cooperation when it is seen as legitimate. Otherwise punishment is treated as mere aggression. Thus punishment is most effective in controlling freeriding only in late-evolving social environments; environments in which something like norms of cooperation are established and salient to all. So the lack of a master model masks an ambiguity about the emergence of punishment. To explain the stability of cooperation, it seems as if it must emerge early, but then its emergence cannot depend on cognitive capacities that probably evolved only in late hominins.

The natural modelling strategy is decompositional: to take each aspect of the cooperation complex (punishment, norm-sensitivity, reputation effects, insider-outsider discrimination, moral emotions) and to model its emergence and stability separately. This strategy makes it easy to tacitly assume, in modelling the evolution of one ingredient of the cooperation stew, backgrounded but important factors that have yet to be cooked. That is one reason why we need a detailed and explicit scenario specifying the changing lineage as a whole, that is, a narrative.

In reconstructing the past, then, narrative and models form a mutually constraining—and supporting—set of epistemic tools. Highly complex explananda like the evolution of human cooperation are resistant to approaches which depend solely on the decomposition and abstraction which enables modellers to probe aspects of constituent dynamics in isolation. For highly complex, multi-factorial, and multi-stage causal trajectories there are no master-models to be had, and so we must instead combine narratives and models, allowing us to navigate between the trade-offs generated by complexity. There is a lesson here for story-tellers as well. Historians often—perhaps characteristically—shy away from abstract, formal approaches to explanation and reconstruction (precisely because these distract from the contingent details). But our discussion suggests that even when historians aim for detailed,

[19] The social world of great apes is not typically one of interactions only amongst close kin, so cooperation evolved in a world of overlapping, only partially coordinated evolutionary interest: kin selection might be some of the story, but it is not most of the story.

narrative explanations, modelling can play an important role in getting there.

5. Conclusion

Brown emphasizes the role of imaginative storytelling in her history of the Lewis Chessmen. Such story-telling has its dangers: when based on idle speculation, the seductive qualities of a good tale can be misleading. However, we have seen that in both history and historical science these worries can be overplayed and can obscure the real benefits—in fact the centrality—of developing narratives in successful historical reconstruction. Speculation is often productive: it furthers our epistemic reach by enabling us to identify the diverse lines of evidence knowledge of the past requires. Further, narratives are themselves embedded in—and part of—that confluence of evidence. And this makes coherence a serious hurdle; the articulation of a narrative is a significant epistemic achievement, and becomes more serious as the historical sciences progress. Finally, narratives do not play second-fiddle to formal, idealized models. The relationship is instead a partnership where each partner compensates for the limits of the other. This last point deserves another: insofar as formal, quantified methods and means of expression are unsuitable to the development of narratives, our argument that narrative is central to historical reconstruction makes a case for not over-playing the importance of the formal and the quantified in legitimate science.

Acknowledgements

This paper was presented at the 2016 Canadian Society for the History and Philosophy of Science conference, we're grateful for the feedback we received. Part of the research was supported by the European Commission, and the Templeton World Charity Foundation (TWCF0128/AB82) (the opinions expressed in this publication are those of the authors and do not necessarily reflect the views of TWCF).

References

Aunger, R. (1995). On Ethnography: Storytelling or science? *Current Anthropology, 36*(1), 97-130.
Beatty, J. (2016). What are narratives good for? *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 58*, 33-40.
Binford, L. (1977). General introduction. In L. Binford (Ed.), *For theory building in archaeology*. New York: Academic Press.
Boehm, C. (2012). *Moral Origins: The Evolution of Virtue, Altruism, and Shame Basic Books*. New York, 2012.
Bowles, S., & Gintis, H. (2011). *A cooperative species: Human reciprocity and its evolution*. Princeton University Press.
Boyd, B. (2009). *On the nature of Stories: Evolution, cognition and fiction*. Cambridge: Harvard University Press.
Brown, N. M. (2015). *Ivory Vikings: The mystery of the most famous chessmen in the world and the woman who made them*. Macmillan.
Cleland, C. E. (2002). Methodological and epistemic differences between historical science and experimental science. *Philosophy of Science, 69*(3), 447-451.
Cleland, C. E. (2011). Prediction and explanation in historical natural science. *The British Journal for the Philosophy of Science, 62*, 551-582.
Currie, A. M. (2014). Narratives, mechanisms and progress in historical science. *Synthese, 191*(6), 1163-1183.
Currie, A. (2015). Marsupial lions and methodological omnivory: Function, success and reconstruction in paleobiology. *Biology & Philosophy, 30*(2), 187-209.
Currie, A. (2016). Hot-blooded Gluttons: Dependency, coherence, and method in the historical sciences. axw005 *The British Journal for the Philosophy of Science*.
Currie, A., & Turner, D. (2016). Introduction: Scientific knowledge of the deep past. *Studies in History and Philosophy of Science, 55*, 43-46.
Davis, D. (1964). The giant panda: a morphological study of evolutionary mechanisms. *Fieldiana Zoology Memoirs, 3*, 1-339.
Farlow, J. O. (1990). Dinosaur energetics and thermal biology. In Weishampel, Dodson, & Osmolska (Eds.), *The dinosauria* (pp. 43-55). Berkely: University of California Press.
Finkelman, L. (2016). *The unspecific unicorn Extinct*. http://www.extinctblog.org/extinct/2016/4/25/what-if-anything-is-a-unicorn
Gächter, S., & Herrmann, B. (2009). Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1518), 791-806.
Gee, H. (2000). *Deep time: Cladistics, the revolution in evolution*. London: Fourth Estate.
Gould, S. J., & Lewontin, R. C. (1979). The spandrels of san marco and the panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London B: Biological Sciences, 205*(1161), 581-598.
Grobstein, P. (2005). Revisiting science in culture: Science as storytelling and story revising. *Journal of Research Practice, 1*(1), 1.
Herrick, C. N. (2004). Objectivity versus narrative coherence: Science, environmental policy, and the US data quality act. *Environmental Science & Policy, 7*(5), 419-433.
Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science, 319*(5868), 1362-1367.
Hull, D. (1975). Central subjects and historical narratives. *History and Theory, 14*(3), 253-274.
Kosso, P. (2001). *Knowing the Past: Philosophical issues of history and archaeology humanity books*.
Lockley, M. G., Xing, L., Kim, J. Y., & Matsukawa, M. (2014). Tracking lower cretaceous dinosaurs in China: A new database for comparison with ichnofaunal data from Korea, the americas, Europe, africa and Australia. *Biological Journal of the Linnean Society, 113*(3), 770-789.
Meisel, Z. F., & Karlawish, J. (2011). Narrative vs evidence-based medicine—and, not or. *JAMA, 306*(18), 2022-2023.
Porr, M. (2015). Beyond animality and humanity; landscape, metaphor and identity". In Coward, Fiona, … (Eds.), *Settlement, society and cognition in human evolution*. Cambridge University Press.
Roth, P. A. (2008). Varieties and vagaries of historical explanation. *Journal of the Philosophy of History, 2*, 214-226.
Rudwick, M. (1972). *The meaning of fossils*. University Press of Chicago.
Sellers, W. I., Margetts, L., Coria, R. A., & Manning, P. L. (2013). March of the Titans: The Locomotor Capabilities of Sauropod Dinosaurs. *PLoS One, 8*(10), e78733.
Sober, E. (1988). *Reconstructing the past: Parsimony, evolution and inference*. Cambrige (MASS) London: MIT Press.
Sterelny, K. (2012). *The evolved apprentice*. MIT Press.
Sterelny, K. (2016). Contingency and history. *Philosophy of Science, 83*(4), 521-539.
Thornton, I. (1996). *Krakatau: The Destruction and Reassembly of an Island Ecosystem*, 1996. Cambridge: Harvard University Press.
Tomasello, M. (2014). *A natural history of human thinking*. Oxford University Press.
Tucker, A. (1998). Unique events: The underdetermination of explanation. *Erkenntnis, 48*(1), 61-83.
Turner, D. (2007). *Making prehistory: Historical science and the scientific realism debate*. Cambridge University Press.
Turner, D. (2009). Beyond detective Work: Empirical testing in paleontology. In Sepkoski, & Ruse (Eds.), *The paleobiological Revolution: Essays on the growth of modern paleontology*. University of Chicago Press.
Weisberg, M. (2013). *Simulation and Similarity: Using models to understand the world*. Oxford University Press.
White, T. D., Lovejoy, C. O., Asfaw, B., Carlson, J. P., & Suwa, G. (2015). Neither chimpanzee nor human, Ardipithecus reveals the surprising ancestry of both. *Proceedings of the National Academy of Sciences, 112*(16), 4877-4884.
Wylie, A. (2011). Critical distance : Stabilising evidential claims in archaeology. In Philip Dawid, William Twining, & Mimi Vasilaki (Eds.), *Evidence, inference and enquiry*. Oxford University Press/British Academy.
Xing, L., Lockley, M. G., Zhang, J., Milner, A. R., Klein, H., Li, D., … (2013). A new Early Cretaceous dinosaur track assemblage and the first definite non-avian theropod swim trackway from China. *Chinese Science Bulletin, 58*(19), 2370-2378.