

Causal properties of narrative explanations tell us what they are and where we'll find them

Margaret Farrell, UC Irvine

Abstract

Some scientists produce explanations that seem to have a story-like, or *narrative* form. Such explanations are especially prominent in the historical sciences, including evolutionary history. To achieve an understanding of explanatory practice in historical science as well as a fuller understanding of the diversity of explanatory practices across the sciences, philosophers of science have proposed accounts of narrative explanation that seek to identify their unique explanatory features. However, such accounts neglect or misidentify the features that distinguish mere narratives from narrative explanations, and narrative explanations from other causal explanations. In this paper, I propose a novel account of narrative explanation that situates it within an interventionist account of causal explanation. Using my account, I propose three dimensions along which narrative explanations can be evaluated and draw several consequences for outstanding disagreements in the literature on narrative explanations. Ultimately, my account clarifies and justifies narrative explanation as a legitimate and distinct form of explanation.

Introduction

Some scientific explanations appear to take the form of a story – especially those in paradigmatically historical sciences like geology, paleontology, and evolutionary biology. Mass extinction events, extreme climate shifts, and the evolution of traits, e.g., the trichromatic eye, are typically explained by reference to the relevant events and phenomena that preceded them (Turner and AboHamad 2023; Currie 2019b). Philosophers of science have often been skeptical of the explanatory power of these stories, beginning with Carl Hempel (1942), who argued that such stories were only sketches of explanations waiting for a scientist to subsume them under a law-like regularity. “Telling stories” is even a pejorative for a scientist – for example, Stephen Jay Gould and Richard Lewontin's infamous accusation that human sociobiologists were only telling “just-so stories” to explain modern human behaviors (Gould and Lewontin 1979; Gould 1978). More recently, philosophers of science, especially philosophers of historical science, have been interested in whether or not such stories are explanatory, and if so, what makes them explanatory. That is, do the stories that historical scientists give count as a distinctive category of scientific explanation? Answering this question would help us to understand explanatory practice in historical science as well as the diversity of explanatory practices across the sciences

Such accounts either locate the explanatory features of narrative in the configuration of a narrative, or in the relationship between an event and its history. I argue that existing configuration and historical accounts fail to identify the distinct features of narrative explanation and I propose a novel account of narrative explanation that situates it within an interventionist account of causal explanation. Using my account, I propose three dimensions along which narrative explanations can be evaluated: change in the likelihood of the outcome given the introduction of proposed events, empirical support for the occurrence of the proposed events, and empirical support for the causal relationships between proposed events. I then draw

several consequences for outstanding disagreements in the literature on narrative explanations. Ultimately, my account clarifies and justifies narrative as a legitimate form of explanation.

The paper proceeds as follows: in §1, I describe what is required for an account of narrative explanation to show that it is indeed a distinct kind of explanation and argue that existing accounts of narrative and narrative explanation have not met those requirements. In §2, I present my account and illustrate it with an example of a narrative explanation of the evolution of human cognition. In §3, I propose some implications that follow from my account, including three dimensions along which we may evaluate narrative explanations. In §4, I use my account to resolve several outstanding questions in the literature on narrative explanation, and in §5 I discuss the distinctions between my account and mechanistic explanation. In §6 I conclude and offer some suggestions for further applications.

§1 Narratives and narrative explanation

To get our bearings for an account of narrative *explanation*, we can set out a few characteristic features of narrative without yet considering how they might explain. “Narrative” may refer to a story simpliciter. But “narrative” may also refer to something more specific, like the *plot* of a story, or it may refer to something more abstract, like the organizing structure of the story, exclusive of other elements like plot, setting, character development, etc. Moreover, narratives have all kinds of functions, epistemic and otherwise: in addition to explanation, they may communicate, entertain, evoke feelings or memories, exemplify, or teach. Narrative is studied in history, literature, and rhetoric, and the extensive literature on narrative and narrative explanation in philosophy and philosophy of science is itself only a portion of the broader literature. Kim M. Hajek points out that many scholars even find it undesirable to circumscribe narrative as a concept; preferring a flexible concept that admits of degree when applied to activities or representations of one kind or another (Hajek 2022). Due to the wide range in which we find use of narrative, accounts of narrative and narrative explanation are quite diverse as well. Because I will defend an account of narrative explanation, I will discuss only the broad characteristics of narrative representation before turning to existing accounts of narrative explanation.

One central feature of narrative as a form of representation is that a narrative presents a set of ordered and related pieces of information – a narrative is not a list (Morgan 2017; 2022; Hull 1975; Roth 2017; 2020; Richards 1992). Moreover, although authors disagree about whether a narrative must be some kind of *temporal* ordering, in any case it cannot be *simply* temporal – a narrative is not a chronicle of time-ordered but otherwise unrelated events. Another feature of a narrative is that its events are organized by some kind of relevance to a *central subject* (Hull 1975). Hull argues that what makes a narrative seem like a cohesive story is that it is about the same entity throughout. The central subject is a historical individual in Hull’s sense, meaning that its identity is maintained through time through some continuity in its constitutive parts. The central subject unifies the narrative – each part of the story is somehow relevant to that subject (Hull 1975). Narratives also characteristically represent some kind of ‘retrospective’ information (Danto 1962; Roth 2020; Ereshefsky and Turner 2020). Accounts vary in the way that they flesh out this feature. On some accounts (including my own) this only means that a narrative

represents a token event and the relevant events which precede it. On others, the notion of retrospective information is developed semantically, as in Arthur Danto and later Paul Roth's notion of *narrative sentences*, which reference past entities such that the sentence could not have meaning until after the event has occurred (Danto 1962; Roth 2020). The foregoing introduction is about as specific as one can get in giving a general definition of "narrative." If we try to get more specific than that, we quickly see that it is a broad concept, with various meanings in different contexts.

My aim is to identify what it is about the narratives that scientists give that makes them explanatory, and under what conditions a scientist is likely to discover or prefer a narrative explanation. I will argue that narrative explanation is best understood as a special case of causal explanation that emerges when the explanandum is a token event that stands in an unstable causal relationship with the first relevant cause in its causal history. I will need to elaborate on the notions of "stability" and "first relevant cause." As I will explain further in §2, causal stability will be construed in James Woodward's (2003) interventionist account of causation and causal explanation and relevance will be construed as *explanatory* relevance determined by the particular explanandum.

There are, broadly, two ways of approaching an account of narrative (and narrative explanation) in science: configuration accounts and causal accounts (Carrier, Mertens, and Reinhardt 2021; Morgan 2022). Configuration accounts initially emerged in resistance to the mainstream consensus of the Deductive-Nomological (DN) account of scientific explanation (e.g., Danto 1962; Richards 1992; Hull 1975; Mink 1966). DN explanations made the law-like regularities an essential part of explanation; narrative scholars maintained that there must be some other explanatory property that did not have to do with regularities. Configuration accounts locate the explanatory feature of a narrative in the holistic structure or organization of the events and relations among them. On Hull's view, for example, there is distinctly explanatory value in seeing the entire narrative arranged into a singular whole (Hull 1975). Mary Morgan, Derek Turner, Marc Ereshefsky, and Kim Sterelny all offer accounts that highlight the explanatory role of a narrative's configuration (Morgan 2017; Ereshefsky and Turner 2020; Sterelny 2021). Morgan argues that narratives explain through the ordering processes that produce them: *colligation* – collecting the information mutually relevant to a topic – and *juxtaposition* – comparing and contrasting the colligated information and generating a question, which the scientist then uses the narrative to answer (Morgan 2017). Thus, the narrative as a whole represents a scientific question, its answer, and the processes that the scientist followed. Turner and Ereshefsky argue, along similar lines as Hull, that a narrative's "ability to integrate reports of events into a single cohesive description" together with the presentation of those events in a correct temporal sequence give a narrative its explanatory force (Ereshefsky and Turner 2020). Kim Sterelny, in a discussion of methodology for constructing a narrative of human evolution, emphasizes the "achievement" in producing a coherent narrative for a complex explanatory target – bringing together many different events over an extended period is so difficult that producing an internally consistent narrative *at all* should have some epistemic weight (Sterelny 2021).

In general, on a configuration account, it is *coherence* and *consistency* that make narratives explanatory. We can think of coherence as the events of the narrative being mutually relevant to one another in some way or another, and consistency as

the property that events in the narrative do not logically contradict one another. We can see this idea of coherence and consistency in Sterelny's description of the achievement of producing a historical narrative, and in Hull's account, and in Ereshefsky & Turner's emphasis on a "single cohesive description." Coherence is roughly the same idea as Hull's shared relevance to some central subject – relevance to the central subject is what makes a narrative coherent.

It seems intuitive that seeing the whole picture of some phenomenon and facts that are somehow relevant to it is in some sense illuminating. But this intuition is somewhat misleading. In the context of scientific explanation, a proposed narrative explanation is a hypothesis. Consistency and coherence are requirements for any scientific hypothesis. In other scientific contexts, these would be considered a fairly low bar. We would not even consider a hypothesis that is not internally consistent in its constitutive claims or the predictions it makes. Sterelny is probably right that it is much more *difficult* to generate a consistent and coherent hypothesis to explain a very complex token event, because the demands for evidence might be diverse and difficult to access (Turner 2005). Nonetheless, internal consistency is a requirement to consider a hypothesis, not reason to believe it. The explanatory force of a narrative, then, does not come from its coherence or consistency. This suggests that the explanatory force of the narrative will (at least implicitly) rely on the explanatory nature of the relations among the elements of the narrative, at least in the context of scientific explanation. Causal accounts are therefore more promising, since they relate narrative explanation to the common scientific practice of causal explanation.

Causal accounts locate the explanatory power of a narrative in the relationship between the explanandum and the relevant events that precede it – they hold that the body of the narrative explains its final event. Informally, the beginning and middle of the story explain the end because those events cause the outcome. Historical narratives are characteristically ordered in time and consist in events and relations between them. Historical narratives are sequences of events causally related to one another (Currie 2014; Ereshefsky and Turner 2020; Beatty and Carrera 2011; Beatty 2017; 2016; Sterelny 2016; Gould 1990). To construct such explanations, scientists empirically investigate phenomena to discover what that sequence was (Currie and Swaim 2021) – they aim to find "the best approximation of the true causal history" (Swaim 2019).

The various causal accounts of narrative explanation agree that a narrative's explanatory power is due, at least in part, to the causal relationships that it represents. They also agree that a narrative is not just a standard causal explanation, however. If narrative explanation is a distinct category of explanation, there must of course be features that distinguish *narrative* explanation from other types of explanation – a requirement articulated nicely by Roth (2017; 2020). Roth argues that an explanation that could just as well be represented non-narratively is not an "essentially narrative explanation" (Roth 2017, 42).¹ For example, the operation of a car engine has a paradigmatically mechanistic explanation. I may verbally describe the story of gasoline entering the tank and being converted to usable energy to move the car, but it is the mechanism doing the explanatory work, not the narrative description. Thus

¹ This distinction appears in another form by Currie (2014), who argues that there are *simple* and *complex* narrative explanations in historical science. The latter correspond to Roth's essentially narrative explanations.

we are looking for the distinctive features of narrative explanation apart from narrative representation, because not every narrative explains in a way related essentially owing to its narrative form. As we saw earlier, narratives are typically invoked to explain *token* events that are not instances of stable generalizations (laws or otherwise). This observation has motivated several accounts to identify what it is about such events that leads to a narrative explanation, as opposed to a standard causal explanation. The historical accounts proposed thus far in the literature, however, have not successfully shown that narrative explanations are in some way distinctive among causal explanations.

One proposed feature is that narratives typically make reference to *alternative possible histories*. For example, John Beatty argues that narratives are branching historical paths connecting events. They are explanatory because they represent relationships of counterfactual dependence between events (Beatty 2016; 2017). They are especially apt at representing the possible alternative paths that history might have taken (Beatty and Carrera 2011). Indeed the idea that narratives represent some kind of counterfactual information appears so salient that Daniel Swaim (2019) identifies this as a *distinctive* feature of narrative explanations. He argues that narratives explain actual outcomes in contrast to alternative histories (Swaim 2019). Swaim takes this to suggest that narratives characteristically deal with *underdetermined* outcomes. But we must be careful – in the historical sciences, alternative histories are only alternative hypotheses. While it may seem intuitive that narratives compare alternative “ways things might have gone,” this is not in principle different from considering one hypothesis in comparison to another. If it is true that narratives are invoked in situations of underdetermination, then, it is not because of the form of a narrative explanation, as underdetermination is itself not unique to historical hypotheses.² If there is something that distinguishes narrative explanation, it must be something else.

Philosophers have identified a number of other candidate properties, either of narrative explanations themselves or the phenomena that they explain. These concepts are *contingency* (Beatty 2017; 2016; Beatty and Carrera 2011), *robustness* (Sterelny 2016), and *peculiarity* (Currie 2019a).³ On Beatty’s account, narrative explanation is not itself a unique category of explanation that we discover, but rather a useful format for representing highly *contingent* histories. He characterizes a narrative as a branching path of causally connected events, where each branch point represents an event that is “contingent *per se*” (Beatty 2016, 36). A contingent *per se* event was, in a strong metaphysical sense, *not yet determined* at the time of its preceding events -- it is an event whose occurrence sets the historical trajectory off down a path when it *could just as well* have gone down a different one. And it is those contingent events that belong in the narrative, because they are the truly puzzling points that explain why some outcome occurred and not another. It is only histories made up of such choice points are that are “worth narrating” (Beatty 2016, 33); that is, the right kind of history to represent in a narrative explanation. “The more forks in the road on the way to the actual outcome, the more points at which history matters,” he

² Underdetermination may instead be due to a lack of available relevant evidence, although whether less evidence is available in the historical sciences is contentious (Turner 2005; Currie 2018).

³ Swaim (2019) also has an additional feature related to mechanistic explanation; for discussion see §5 on mechanism.

writes with Carrera, noting that “[a] narrative is just right for emphasizing the critical intermediary points, and would be superfluous without them” (Beatty and Carrera 2011, 491–93).

The claim of indeterminacy, however, is a hard pill to swallow for many scientists. And, if nothing else, it is certainly the case that scientists who reject metaphysical indeterminacy construct and put forth narrative hypotheses, and presumably they are doing so on some other basis than locating the points of metaphysical indeterminacy. So even if one is not committed to determinism, it is useful for understanding the explanatory practices of scientists to have an account of narrative explanation that does not depend on such choice points. But notice that without this notion of contingency as indeterminacy, we lose the force of Beatty's account of narrative explanation as distinctive, because we are left only with causal dependence, and this does not help us distinguish which sequences of causal dependence are, or ought to be, represented as narratives. Eric Desjardins gives a formulation of Beatty's account that does not include metaphysical contingency but in which he argues that historicity involves multiple possible past states, multiple possible outcomes, and causal dependence (Desjardins 2011; see also Beatty and Desjardins 2009). Desjardins emphasizes that the path a narrative takes is sensitive to initial conditions (Desjardins 2011). But Turner and Ereshefsky point out that on Beatty's and Desjardins' account, covering-law explanations would count as narrative explanations, because the occurrence of an upstream cause, e.g., placing a piece of copper in sulfuric acid, may be contingent *per se* (we may not have done so), and the occurrence of an effect, e.g., the dissolution of the piece of copper, is counterfactually contingent upon the upstream event (Ereshefsky and Turner 2020, 50). Thus, it seems that this account has not identified what is distinctive about narrative explanations over and above their status as causal explanations.

Adrian Currie has developed an account of narrative explanation in the context of the historical sciences in which he seeks additional properties or features that make narratives distinctive among causal explanations. His account, however, has a similar issue with identifying the properties that distinguish narrative explanations from other causal explanations. Currie's early view on narrative explanation picks out what he calls "embeddedness" and "detail" (Currie 2014). "Detail is a measure of the specificity, complexity and diffusion of the explanans required for explanatory adequacy" while embeddedness describes those explananda that are "accounted for as a token of a type; an instance of a regularity" (Currie 2014, 1169). These two properties lead Currie to distinguish between "simple" and "complex" narratives – complex narratives are high in detail and not embedded, while simple narratives are low in detail and embedded (Currie 2014, 1169). But note here that Currie's simple narratives then have the same trouble that Ereshefsky & Turner pointed out in Beatty's work: paradigmatic covering-law explanations count as simple narratives. Complex narratives are characterized only descriptively: they have high detail and their explananda are not tokens of type-level regularities, but it is not clear why this is the case. In more recent work, Currie introduces a concept that he argues underlies the properties we find in complex narratives: *peculiarity* (Currie 2019a). Peculiarity is a feature of an explanandum (in Currie's account, called a "target"). Certain targets are more peculiar than others, and peculiar targets require narrative explanation. Peculiarity is, in Currie's words, "the extent to which [a target's] modal profile depends on the processes that generate and maintain it"

(Currie 2019, 42). Targets, in Currie's words, “arise from, and are maintained by, processes,” and a target's modal profile is “the likelihood of a target's properties staying the same or changing” combined with “the ways a target may transform under various conditions” resulting in “a map of the various conditions under which a target might change” (Currie 2019a, 42). He then writes, “When peculiarity holds, the target’s modal profile is sensitive to some features of these processes” (Currie 2019a, 42). But note that if “processes” are just those things that generate and maintain targets, then the properties of some target (along with all its potential properties, or potential for change) always depend on the processes that produce them. The degree of peculiarity is supposed to determine whether a target requires a narrative explanation, but there does not seem to be a dimension along which peculiarity could vary. Thus, Currie's account, like Beatty's, boils down to a causal dependence claim. But it is not distinct from other forms of causal explanation and so it cannot pick out distinctive features of narrative explanation.

The final account I discuss comes from Kim Sterelny. He seeks to identify what it is about certain events in history that leads historians to explain them with either a “robust process explanation” or an “actual-sequence explanation,” a distinction he develops following Jackson and Pettit (1992) (Sterelny 2016, 523). Sterelny argues that some historical events can be explained by robust processes while others require a detailed causal sequence leading up to the occurrence of the event. Similar to Desjardins’ account, on Sterelny's view, to be contingent is to be sensitive to small perturbations in some initial condition. For example, Sterelny argues that the first World War was *not* contingent, because there were many conditions in place such that the specific nature of the triggering event (the assassination of Franz Ferdinand) did not make much of a difference (Sterelny 2016, 527–29). WWI was therefore the result of a robust process – it was going to happen regardless of the triggering event that, in fact, began the war. In contrast, the outcome of some battle is contingent, e.g., “If a battle is lost because the admiral in command has a stroke at the crucial moment... [I]ts causes are not likely to be captured by an analysis of the relevant military factors before the battle” (Sterelny 2016, 522). Here, an actual-sequence explanation – a detailed causal story – is required because the occurrence of the event is not very robust – that is, whether or not it would happen in many possible worlds – is comparatively low. The decisions of particular individuals often require such explanations. This is the reason that some larger-scale historical events end up contingent, because human societies developed “command-and-control systems that put individuals in power such that their idiosyncratic decisions, rather than the more robust aggregate behaviors of groups, end up with more causal influence (Sterelny 2016, 534–37). In human history, events that are the result of individuals’ decisions may require actual-sequence explanations because those individual decisions are much more fickle than the aggregate behaviors of groups.

But Sterelny’s explanation types are tied to the *occurrence of the event* – they answer the question of whether, in a global sense, the event *would have happened anyway* under somewhat different circumstances. Accordingly, the examples of events that are explained by robust processes answer this question. But they do so by pointing out that the events in question were causally overdetermined. Sterelny's examples of robust process explanations – WWI, the Nuer conquest in Sudan, the retreat of the Bubonic plague in western and central Europe, the European demographic transition

to smaller family size -- are all examples of causal overdetermination. For each case, Sterelny cites a number of causes that overwhelm the possibilities of history turning out otherwise, such that the particular events in each case cease to make much of a difference taken individually. Because the need for an actual-sequence explanation depends only upon whether or not the event would occur in other possible histories, Sterelny's account leaves open the possibility of an event whose occurrence is highly robust but for which there is no robust process to explain it. The robust processes described in Sterelny's examples are cases of causal overdetermination which, in a sense, *amount* to a robust process: situations of various kinds of instability, conflict, and anticipation of conflict lead reliably to wars. The many causes that make WWI's occurrence robust are *related* to one another – there are similarities in these causes that make the situation overall an instance of a generalization, a token of a type. But it is consistent with Sterelny's account that a number of causal histories, unrelated to one another (except for the fact that they produce the same outcome), could simultaneously occur. Suppose that just as my cat begins to push a glass vase off the edge of my kitchen table, my partner bumps into the table as he walks by, and at the same time, an earthquake strikes, shaking the table and causing the vase to fall off. The event is counterfactually robust, but it could not be explained by a robust process. Sterelny's account would tell us we ought to find a single robust process, where instead what explains why the event happened or not is the occurrence of three convergent histories. Therefore, while actual-sequence explanation tracks with a narrative explanation, Sterelny has not identified under what conditions we need or should expect to find one. Sensitivity to initial conditions will exclude events that are counterfactually robust but not instances of robust generalizations.

Because narratives explain token events that are not simply instances of generalizations, something like contingency or sensitivity to initial conditions or path dependence likely *is* at the root of what calls for a narrative explanation. What each of these accounts fails to do is distinguish narrative explanations from other simply causal explanations. We might conclude that therefore narrative explanations are nothing more than causal explanations, without any special formal features – perhaps this means it is only their content that makes them narrative or historical explanations, or that their distinctive features are not in their explanatory form but in their use as rhetorical or communicative tools. I do not think that we ought to take this road, despite my criticisms of these authors, because I think we can say something slightly more nuanced, without giving up that narrative is a useful category of explanation (even though I will argue that it is a useful *subcategory* of *causal explanation*). The persistent challenge is to find what it is about a narrative explanation that not every causal explanation has, or, in Sterelny's case, to do so in an appropriately specific way. I can now set out the task for my account much more clearly: what is it about an event of interest that leads us to explain it with a narrative? Is it a feature of the event itself, of its causal relationships with prior events, or a particular set of explanatory aims? The answer, I argue, is a combination of the three: explanatory aims determine the first relevant cause of a token explanandum event, and when that causal relationship is unstable, a narrative is required to fill in the space between. My argument for this claim will involve making good on the notion of a first relevant cause, as well as a concept of causal stability that I promised earlier. In the next section, I argue that a narrative explanation is the causal sequence that links a token event to its first relevant cause when that overall

causal relationship is *unstable*. Instability is construed as in James Woodward's interventionist account of causation, as are the rest of the causal explanatory properties of narratives. Ultimately, my account says that it is not entirely accurate to speak of conditions that *require* narratives as if we have a choice of the type of explanation we might use; rather there exist causal sequences, that, as a result of their causal properties and our explanatory aims, *have* the properties that characterize narratives: central subjects, retrospective representation, coherence and consistency, etc. Thus, when we represent, relay, tell that causal sequence in response to a why-question, we are giving a narrative explanation.

§2 Narrative explanations and causal instability

Because not every narrative explains, it is a mistake to begin an account of narrative explanation by asking how the characteristic features of narrative representation might be explanatory. I approach the problem instead by asking what conditions could produce an explanation that resembles a narrative. I argue that under certain conditions, the *causal* explanation of a target of interest takes on the characteristic features of a narrative, and this is what a narrative explanation is. My account is summarized as follows:

A narrative explanation is the causal explanation that results when the explanatory target is a token event that stands in an unstable relationship with the first relevant cause in its causal history.

In other words, a narrative explanation is a causal explanation with narrative properties. I will now illustrate this account with an example from human cognitive evolution. I use this example because explanations that resemble narratives in evolutionary history are particularly controversial – indeed, “storytelling” is a commonplace pejorative for explanations in this field (stemming from critiques by Stephen J. Gould and Richard Lewontin (1978; 1979)). By examining this example, we will see how narrative explanation works and that there is nothing inherently suspicious about an explanation that sounds like a story. The case I will examine is Sarah Blaffer Hrdy’s narrative hypothesis for the evolution of uniquely human cognition, as presented in her book *Mothers and Others* (Hrdy 2011). While this hypothesis is ultimately incomplete,⁴ it has been accepted as identifying an important component of hominin⁵ evolutionary history, and it is also an illustrative example of a genuinely narrative explanation.

Figure 1 shows a graphic summary of Hrdy’s narrative. The end of Hrdy’s narrative is the explanatory target: the emergence of the particular set of species-typical cognitive capacities of modern humans. These include capacities for exceptionally prosocial and cooperative behaviors, linguistic capacities, and complex theory of mind. Hrdy also agrees with the consensus view that the evolution of these capacities in hominin brains was due in part to an increase in brain size. The beginning of Hrdy’s narrative is an earlier event: the emergence of a species with a

⁴ Theorists typically accept that the emergence of cooperative breeding was an important event in the evolution of human cognition but contend that other factors were equally or more important.

⁵ “Hominin” refers to humans and all of the extinct species with whom we share more recent common ancestors than the LCA that we share with chimpanzees & bonobos.

particular set of cognitive capacities, ancestral to modern humans and our closest extant relatives, chimpanzees and bonobos. This last common ancestor (LCA) could, among other things, engage with and imitate others at least in infancy and had some theory of mind capacities. The beginning of a narrative is the hypothesized *first relevant cause*. The first relevant cause is determined by the explanatory aims we set out at the start of inquiry and the background theory that informs the explanatory contrast of interest. Here, the explanatory contrast of interest is why modern humans developed the set of cognitive traits that we have, rather than some other possible set. Background evolutionary theory tells us that humans share ancestors with modern species that have a wide range of different sets of cognitive traits, and that the relevant *uniquely human* traits are those possessed by us, and not by our closest evolutionary relatives or the most recent common ancestor that we share with them. Phylogenetic history tells us that changes that determined the unique set of cognitive capacities of modern humans occurred during the period of time between the emergence of the LCA species and the present (more specifically, sometime during the Late Miocene and Pleistocene). The challenge for theorists like Hrdy is to figure

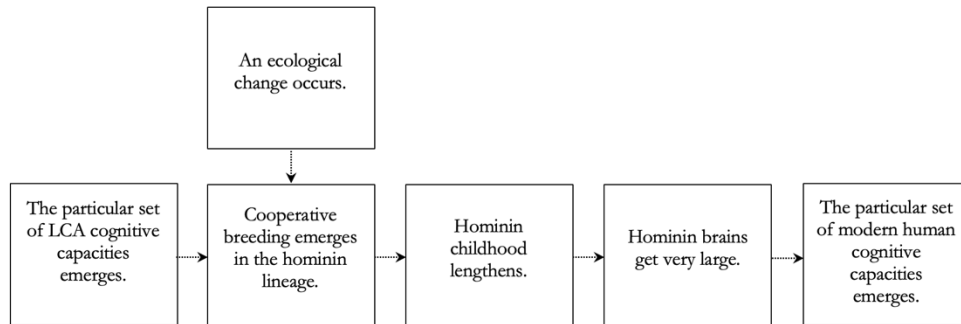


Figure 1. Summary diagram of Hrdy's hypothesis for the evolution of uniquely human cognition.

out why the hominin lineage ended up with our particular set of cognitive capacities and not another set.

The middle of the narrative consists of several major events that connect the first relevant cause to the explanatory target. Hrdy's primary proposal is the emergence of cooperative breeding in the hominin lineage, though she first postulates a less specific event – some ecological change that introduced the selective pressures that led to the emergence of cooperative breeding. The rest of the events are consequences of this important development, in combination with selective pressures that were largely already present but unable to affect the hominin lineage prior to the emergence of cooperative breeding. Cooperative breeding allows offspring to spend longer developing in childhood. With more time and energy to devote to development, brain size could then increase (given that selective pressures for larger brains were already present). Then, with larger brains to work with, selective pressures began to shape the cognitive capacities that we now observe in modern humans. Hrdy's explanation has those properties that characterize narrative representation. She describes a set of events, ordered in time, ending with a token event. Its central subject is the hominin lineage. But Hrdy's narrative is clearly

presented as an *explanation* for the emergence of uniquely human cognitive capacities. Why does the resulting explanation have the properties of narrative representation?

Hrды wants to explain why our particular set of cognitive capacities evolved, rather than some other set. She poses this question as “why us and not them?” (Hrды 2011, 63) – that is, why did the hominin lineage acquire, e.g., complex theory of mind, while other lineages that share the LCA did not? This is the relevant *contrastive focus* for the explanatory target – it constrains the many possible ways of answering why human cognition is the way it is to the causal factors that make the difference between the hominin lineage and its neighboring lineages.⁶ Hrды does not give her story in the form of a precise causal graph, but we can do so easily. The explanatory target and the emergence of the LCA can be represented as causal variables: B for the beginning set of cognitive capacities, and E for the end, with, say, b the particular set of the LCA and e the particular set of modern humans. Note that before Hrды even begins to fill in the narrative, there is a causal relationship here. Adopting a minimal criterion for causation from Woodward (2003), the relationship between B and E is causal:

X causes Y if and only if there are background circumstances R^7 such that if some (single) intervention that changes the value of X (and no other variable) were to occur in R , then Y or the probability distribution of Y would change (Woodward 2010, 63).

The causal relationship $B \rightarrow E$ is borderline trivial; the traits of an ancestral species are obviously causally related to the traits of a descendent species. And in fact other accounts of narrative and narrative explanation have identified counterfactual dependence as an important feature that drive their explanatory force (Beatty 2016; 2017). But observing that construction of the narrative begins with a hypothesized causal relationship between the beginning and the end is crucial to clarifying its distinctive properties. The relationship $B \rightarrow E$ on its own only holds in very few sets of background conditions. The set of cognitive capacities of an ancestral species fully specifies the set of cognitive capacities of a descendent species only in circumstances in which *no evolutionary change occurs* – that is, when $B = E$. Thus, even though a minimal causal-explanatory view suggests that B explains E , it is trivially explanatory at this point. Nevertheless, our choice of explanatory target and our background evolutionary theory dictates that our story starts here, with the cognitive capacities of the LCA, and ends with the cognitive capacities of modern humans. This property of a causal relationship holding only under very few sets of background conditions is called “instability” in the interventionist framework (Woodward 2010). If the causal relationship $B \rightarrow E$ were very stable, and we had evidence for the occurrence of b and e , then the causal explanation might be adequate. But it is not so stable. The relationship can be disrupted by all kinds of background conditions; in fact, the same b did lead to different e_i – chimpanzees and bonobos witness this. This observation allows us to make sense of what theorists do next, which is to *fill in the space* between

⁶ My use of the term “contrastive focus” comes from Lauren Ross (2019), who argues that setting this focus is a necessary part of fully specifying the explanatory target of interest.

⁷ I have used “R” rather than Woodward’s “B” to avoid confusion with my use of “B” for the first causal variable in Hrды’s narrative.

the beginning and the end. Each time that Hrdy proposes an *event* in the narrative, she is at the same time introducing a causal variable into her explanation. The rest of Figure 1 can be straightforwardly represented by a causal graph in Figure 2:

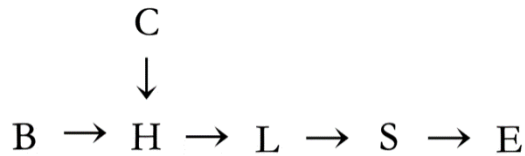


Figure 2. Causal graph of Hrdy’s narrative explanation. B = the set of LCA cognitive capacities, C = the occurrence of a particular ecological change, H = the emergence of cooperative breeding in the hominin lineage, and so on, corresponding to Figure 1.

The causal graph in Figure 2 represents the narrative that Hrdy proposes and illustrates that it is a legitimate candidate causal explanation. But it is the instability of the causal relationship $B \rightarrow E$ that demanded additional causal steps be added. As Hrdy imports features that were background conditions relative to $B \rightarrow E$, the causal structure becomes relatively more stable.

While the addition of causal variables increases the stability of the causal system relative to $B \rightarrow E$, the overall causal structure of the explanation need not be stable in some absolute sense; we only must be confident that the story occurred as described. And indeed, when Hrdy builds the narrative, she does not proceed primarily by hypothesizing causal variables to introduce. Narrative theorists add *events* – values of the variables that we represent in the causal graph. The introduction of an event should increase the probability of the particular value e conditioned on the events included in the narrative. That is, when we begin with the unstable causal relationship $B \rightarrow E$, the conditional probability $\Pr(e|b)$ is different from $\Pr(e)$, but the difference is almost negligible. Each additional event should increase the probability of e conditioned on the events of the narrative.

To summarize: a narrative explanation is a causal explanation with narrative properties, which we find in situations in which the explanatory target and its beginning stand in an unstable causal relationship. Narrative construction proceeds by the introduction of events whose occurrence increases the probability of the end value conditioned on the events of the narrative; as the relevant causal variables are added, the causal system is increasing (however slightly) in overall stability. A narrative explanation is not chosen from among other possible explanatory strategies,⁸ rather, it is the form that a causal explanation takes under certain conditions.

§3 Evaluating narrative explanations

This causal interpretation allows us to evaluate narrative explanations along three dimensions: change in the likelihood of the outcome given the introduction of

⁸ That narrative explanation is not always a pragmatic choice is also a feature of Currie’s (2014) account.

proposed events, empirical support for the occurrence of the proposed events, and empirical support for the causal relationships between proposed events.

Causal Graphs:



Narrative Diagrams:

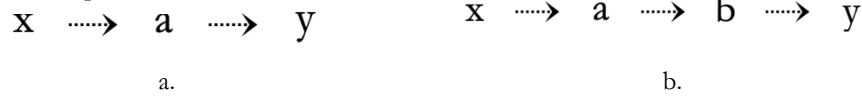


Figure 3. Causal graphs and narrative diagrams for two narrative explanations. As in Figs. 1 & 2, causal graphs are represented with solid arrow and capital letters for variables while narrative diagrams are represented with dotted arrows and lowercase letters for instances.

The difference between Figure 3a and 3b captures a comparison of the first dimension. Often, but not always, the introduction of additional events will increase the conditional probability of the outcome – this is why often, a narrative with “more detail” is preferable to a less detailed one. The causal explanation in Figure 3b contains more causal variables and more proposed events. Assuming that all the causal relationships proposed and the occurrences of x , a , b , etc. are well-supported, 3b would be the better explanation.

This helps make sense of something puzzling in several discussions of narrative explanation. Currie (2014) observes that narrative explanations have more detail than other explanations. He states that “detail” is a “measure of the specificity, complexity, and diffusion of the explanans” (Currie 2014, 1168). But Currie leaves unanswered just how these measures apply to an explanans, and more importantly, just why we should expect a narrative explanation to have more detail. Marc Ereshefsky and Derek Turner (2020) similarly argue that a “thicker” narrative (one which contains more causally relevant events) is better than a “thinner” one, all else equal. They argue that a thicker narrative brings the historical path to the outcome “into sharper focus” (Ereshefsky and Turner 2020, 53). But this metaphor only begs the question; to bring something into sharper focus is only to increase resolution. It does not tell us *why* higher resolution is better. Recall from §1 that the motivation for configuration accounts of narrative: that producing a narrative alone is producing an explanation of some kind, and especially the emphasis that Sterelny placed upon the achievement of a narrative and how this speaks in favor of its truth. Why should the internal consistency of a long, complicated narrative speak in its favor? What Currie, Ereshefsky, Turner, and Sterelny are actually latching on to is the idea that just producing the narrative is an epistemic achievement of figuring out how to increase the conditional probability of the particular outcome that we observe in modern humans.

The second two dimensions for evaluating narrative explanation concern empirical support. In Figure 3a, I have represented both the causal graph and the chain of events for a short narrative. In order to establish the adequacy of this explanation, we need evidence that the *causal relationships* $X \rightarrow A$ and $A \rightarrow Y$ hold, *and* evidence that events x , a , and y occurred. Surely some evidence may play both roles, but the two can come apart. Hrđy’s evidence that cooperative breeding causes

an increase in tolerance for food-sharing is drawn from studies on various extant monkey species (Hrady 2011, 92–98). But separate evidence is needed to claim that the emergence of cooperative breeding and an increase in tolerance of food-sharing occurred in the hominin lineage. This separation captures an intuition that many philosophers hold about narratives: there is a way for a narrative explanation to be somehow epistemically *good* before we know whether or not it is *true*. If there is a substantial amount of empirical support for the causal relationships $X \rightarrow A \rightarrow Y$, the narrative has empirical support, but not yet support for the occurrence of particular values x , a , and y .

§4 Implications of a causal interpretation

My account resolves several disagreements in the literature on narrative and historical explanation. First, notice that Hrady’s narrative contains events that happen to the hominin lineage, and ‘external’ events like ecological changes that are causally relevant to events that happen to the lineage. From this we can state a causal interpretation of how a central subject organizes a narrative: causal variables either represent events that reference the central subject or are causally relevant to events that reference the central subject. Second, the question of whether retrospective knowledge is essential to narratives and what impact this has on the way that they explain is clear in a causal framework: that a narrative explanation is retrospective is derivative of the fact that it causally explains a token event. Third, the broad scope of a causal framework like interventionism means that there is no restriction on the kind of cause that is included in the narrative explanation (whether it is an event or some regularity that becomes causally relevant). This resolves questions in the narrative explanation literature about the status of regularities in historical explanations (Currie 2014; 2018; Jeffares 2008; Kosso 2001; Kranke 2022; Turner 2009). Nothing in the form of a narrative explanation dictates whether the causes involved are part of regularities.

The two dimensions of evidence for narratives that I propose make sense of why philosophers are sometimes tempted to treat proposed narrative explanations as “how-possibly explanations” (e.g., Kitcher (2011)). Others have articulated that a how-possibly explanation only shows that some outcome is *possible* under certain conditions in contrast to *impossible* (Brandon 1990; Reydon 2012). Proposed narratives that explain why one event happened in contrast to another do not do this. They are proposed hypotheses for the *actual* course of history. The temptation to call these a kind of *explanation* makes sense once we see how a narrative explanation can have empirical support for the truth of the causal claims it makes *without* (yet) having empirical support for the occurrence of the events that instantiate those causal claims. Producing an internally consistent narrative is an epistemic achievement, but not an explanatory one.

Finally, the broad scope of a causal framework like interventionism means that there is no restriction on the kind of cause that is included in the narrative explanation (whether it is an event or some regularity that becomes causally relevant). This resolves questions in the narrative explanation literature about the status of regularities in historical explanations (Currie 2014; 2018; Jeffares 2008; Kosso 2001; Kranke 2022; Turner 2009). Nothing in the form of a narrative explanation dictates whether the causes involved are part of regularities.

§5 Narratives and mechanisms

An important question remains: Why is a narrative not a mechanism? Mechanistic explanation is a prominent type of explanation in science, especially in the life sciences (see, e.g., Machamer, Darden, and Craver (2000)). There are several accounts that characterize narrative explanation not as a distinctive category within causal explanations more broadly, but as a certain kind of mechanistic explanation. There have been two primary strategies for doing so: the first is to characterize narratives and/or histories as a kind of one-off mechanism. This is the strategy taken by Stuart Glennan (2010) and Daniel Swaim (2019). The other is to characterize a historical explanation as a string of mechanistic explanations ordered in time, an approach developed by Brett Calcott (2009).

Glennan proposes that narratives and histories are “ephemeral mechanisms” — they have all the properties of mechanisms, except they only occur once. The relationships among the ‘entities’ are stable and change-relating, but the entities do not come together very often, so the mechanism has only had one chance to run. He writes,

Specifically, I take an ephemeral mechanism to be a collection of interacting parts where:

1. the interactions between parts can be characterized by direct, invariant, change-relating generalizations
2. the configuration of parts may be the product of chance or exogenous factors
3. the configuration of parts is short-lived and non-stable, and is not an instance of a multiply-realized type (Glennan 2010, 260).

Similarly, Swaim introduces the idea of a “mechanistic setup” (Swaim 2019, 7). Both accounts seek to legitimize the causal explanatory nature of a narrative by showing that it is characterized by all the same features as a mechanism except for the overall regularity of the mechanism. But this does not apply to, for example, Hrdy’s narrative. Some of the causal relationships in Hrdy are not invariant, and some are, but the *overall* relationship between beginning and end is not. The causal relationships within the history may be any degree of invariant, and the pieces of the narrative do not have straightforward interpretations as entities, because they are events. But without the regularity of a mechanism’s overall operation and the stable change-relating relationship between entities, it is neither descriptively accurate to call a narrative a mechanism nor explanatorily helpful, because the regularity of the operation of a mechanism is integral to its status as an explanation.

Brett Calcott follows a different line of argument. He introduces the notion of a *lineage explanation*, with the aim not to dissolve the idea of distinctly narrative explanations, but rather to be descriptively accurate to the explanations generated in evolutionary developmental biology. Some authors of human cognitive evolution narratives even take their ideal or aim to be something like Calcott’s lineage explanation (e.g., Sterelny and Planer (2021, 5)). Lineage explanations trace how a mechanism could change over time to show continuity between one form and another. In the evolutionary-developmental explanation of the evolution of the

focused lens eye, he writes, “the same three layers of cells are present throughout the stages. Their shape, size and relation to one another are modified, but the credibility of continuity between the stages is justified because there is no radical change in parts, or their properties, during the transitions (Calcott 2009, 64).” This example illustrates the two-dimensional nature of a lineage explanation: a production dimension and a continuity dimension. The production dimension captures that each stage of the explanation is an independent mechanistic explanation of how a character is produced. Crucially, what it is that links the stages is not important. He also argues that natural selection does not play a significant role, if any, in a lineage explanation of a trait:

Likewise, in a lineage explanation, the fact that natural selection caused the change is not part of the explanation. What matters is that we describe the mechanism in a way that we can see how a small modification to some component could have, by whatever process, enabled such a change in the mechanism (Calcott 2009, 67).

But narratives like Hrdy’s about are not lineage explanations. A mechanistic explanation at each stage of hominin cognitive evolution is not even desirable, leaving aside the empirical challenge of ever accessing that kind of information. Further, in Calcott’s lineage explanations, the continuity of the sequence of stages only comes from their being *close enough* to one another – “the credibility of continuity between the stages is justified because there is *no radical change* in parts, or their properties, during the transitions” (Calcott 2009, 64, emphasis mine). But here Calcott is relying on an unspecified sense of “radical change.” In the case of the evolutionary-developmental explanation of the eye, this sense of “radical change” is probably based on a theorist’s trained sense of plausibility relating to mechanical complexity, evolvability, number or kind of mutation that could produce the posited change, etc. This is effectively an informal appeal to legitimate background theory. In contrast to Calcott’s characterization, the judgment that one stage is close enough to the next is crucial to the explanation overall. In an evolutionary-developmental explanation of the mammalian eye, informal judgments that a change is small enough may be harmless, but only because they happen to be good proxies for judgments justified by background theory.

But in the context of human cognitive evolution, these evolutionary reconstructions will indeed be stories of natural selection – changes in selective regimes will often form the link between one stage to the next. In this context it is much easier to see that the transitions from one stage to the next must be empirically supported, not idiosyncratic judgements of whether a difference is small enough. Thus, Calcott’s account does not accommodate the features of historical narratives, and so cannot account for their unique explanatory features. Further, given that the judgments of “no radical change” in the evolutionary-development case are in fact proxies for judgments supported by legitimate background theory, it is possible that Calcott’s account does not capture all such evolutionary-development explanations either. They may instead be narrative explanations in the sense I argue for in this paper. In any case, neither Glennan and Swaim nor Calcott’s accounts of mechanism can account for narrative explanation.

§6 So, just storytelling? Conclusions and applications

On my account, narrative explanations are causal explanations that result from a set of conditions involving both our explanatory aims and the causal relationships rendered relevant by those explanatory aims. The explanatory value does not arise from the narrative form; rather the form of the narrative emerges from the causal properties of the event we are interested in explaining. My account situates narrative explanation within a well-established framework of scientific explanation, clarifying the explanatory form and justifying its power. It also, as I have shown, resolves several ambiguities in the literature on narrative explanation.

I conclude with a note on applications of my account. One especially salient context for narrative explanation, as we have seen, is evolutionary history. Despite the past criticisms of storytelling, a number of recent, well-received hypotheses for the evolution of uniquely human cognition which are presented as narratives – not only Hrdy’s narrative as discussed here, but in hypotheses from Michael Tomasello, Celia Heyes, Kyle Stanford, Kim Sterelny, and Ronald Planer.⁹ Why should these theorists choose to present narratives? My account shows that there is nothing suspect about these explanations in virtue of their narrative form alone.¹⁰ Producing a story is not just storytelling; it is the construction of a legitimate causal explanation demanded by the properties of the causal history of the explanatory target. My account may even distinguish these hypotheses as an altogether different approach to the evolution of human cognition compared to alternatives like Sociobiology and Evolutionary Psychology. Further, because of its evaluative power, my account may offer a framework in which to understand other features of this approach. A more precise understanding of narrative explanation will therefore be invaluable to making progress in paradigmatically narrative fields of inquiry like evolutionary history.

References

- Beatty, John. 2016. “What Are Narratives Good For?” *Studies in History and Philosophy of Biological and Biomedical Sciences* 58 (August):33–40.
<https://doi.org/10.1016/j.shpsc.2015.12.016>.
- . 2017. “Narrative Possibility and Narrative Explanation.” *Studies in History and Philosophy of Science Part A* 62 (April):31–41.
<https://doi.org/10.1016/j.shpsa.2017.03.001>.
- Beatty, John, and Isabel Carrera. 2011. “When What Had to Happen Was Not Bound to Happen: History, Chance, Narrative, Evolution*.” *Journal of the Philosophy of History* 5 (3): 471–95.
<https://doi.org/10.1163/187226311X599916>.
- Beatty, John, and Eric Cyr Desjardins. 2009. “Natural Selection and History.” *Biology and Philosophy* 24 (2): 231–46.
- Brandon, Robert N. 1990. *Adaptation and Environment*. Princeton: Princeton University Press.

⁹ See Heyes (2018), Sterelny (2012; 2021), Sterelny & Planer (2021), Stanford (2017), Tomasello (2014; 2016; 2019).

¹⁰ See also Olmos (2022). Other authors have responded to just-so story criticisms by pointing out other non-explanatory epistemic functions, e.g., Lennox (1991).

- Calcott, Brett. 2009. "Lineage Explanations: Explaining How Biological Mechanisms Change." *British Journal for the Philosophy of Science* 60 (1): 51–78.
<https://doi.org/10.1093/bjps/axn047>.
- Carrier, Martin, Rebecca Mertens, and Carsten Reinhardt, eds. 2021. *Narratives and Comparisons: Adversaries or Allies in Understanding Science?* transcript Verlag.
<https://doi.org/10.2307/j.ctv2f9xs9f>.
- Currie, Adrian. 2014. "Narratives, Mechanisms and Progress in Historical Science." *Synthese* 191 (6): 1163–83. <https://doi.org/10.1007/s11229-013-0317-x>.
- . 2018. *Rock, Bone, and Ruin*. MIT Press.
- . 2019a. "Scientific Knowledge of the Deep Past: History Matters." *Elements in the Philosophy of Science*.
- . 2019b. "Simplicity, One-Shot Hypotheses and Paleobiological Explanation." *History and Philosophy of the Life Sciences* 41 (1): 10.
<https://doi.org/10.1007/s40656-019-0247-0>.
- Currie, Adrian, and Daniel Swaim. 2021. "Past Facts and the Nature of History." *Journal of the Philosophy of History* 16 (2): 179–206.
<https://doi.org/10.1163/18722636-12341457>.
- Danto, Arthur C. 1962. "Narrative Sentences." *History and Theory* 2 (2): 146–79.
<https://www.jstor.org/stable/2504460>.
- Desjardins, Eric. 2011. "Historicity and Experimental Evolution." *Biology and Philosophy* 26 (3): 339–64.
- Ereshefsky, Marc, and Derek Turner. 2020. "Historicity and Explanation." *Studies in History and Philosophy of Science Part A* 80 (April):47–55.
<https://doi.org/10.1016/J.SHPSA.2019.02.002>.
- Glennan, Stuart. 2010. "Ephemeral Mechanisms and Historical Explanation." *Erkenntnis* 72 (2): 251–66. <https://doi.org/10.1007/s10670-009-9203-9>.
- Gould, Stephen Jay. 1978. "Sociobiology: The Art of Storytelling." *New Scientist* 80 (1129): 530–33.
- . 1990. *Wonderful Life: The Burgess Shale and the Nature of History*. W. W. Norton & Company.
- Gould, Stephen Jay, and Richard C. Lewontin. 1979. "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme." *Proceedings of the Royal Society of London - Biological Sciences* 205 (1161): 581–98.
<https://doi.org/10.1098/rspb.1979.0086>.
- Hajek, Kim M. 2022. "What Is Narrative in Narrative Science? The Narrative Science Approach." In *Narrative Science*, edited by Mary S. Morgan, Kim M. Hajek, and Dominic J. Berry, 1st ed., 31–58. Cambridge University Press.
<https://doi.org/10.1017/9781009004329.003>.
- Hempel, Carl G. 1942. "The Function of General Laws in History." *The Journal of Philosophy* 39 (2): 35–48.
- Heyes, Cecilia. 2018. *Cognitive Gadgets: The Cultural Evolution of Thinking*. Cambridge, MA: Belknap Press.
- Hrdy, Sarah Blaffer. 2011. *Mothers and Others: The Evolutionary Origins of Mutual Understanding*. The Belknap Press.
- Hull, David L. 1975. "Central Subjects and Historical Narratives." *History and Theory* 14 (3): 253–74. <https://www.jstor.org/stable/2504>.

- Jackson, Frank, and Philip Pettit. 1992. "In Defense of Explanatory Ecumenism." *Economics & Philosophy* 8 (1): 1–21.
<https://doi.org/10.1017/S0266267100000468>.
- Jeffares, Ben. 2008. "Testing Times: Regularities in the Historical Sciences." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 39 (4): 469–75.
<https://doi.org/10.1016/j.shpsc.2008.09.003>.
- Kitcher, Philip. 2011. *The Ethical Project*. Harvard University Press.
- Kosso, Peter. 2001. *Knowing the Past: Philosophical Issues of History and Archaeology*. Amherst: Humanity Books.
<https://www.cambridge.org/core/journals/american-antiquity/article/abs/knowning-the-past-philosophical-issues-of-history-and-archaeology-peter-kosso-2001-humanity-books-amherst-new-york-200-pp-isbn-1573929077-4500-cloth/952F5F1199692BD0ADD69413AF43A61F>.
- Kranke, Nina. 2022. "Two Kinds of Historical Explanation in Evolutionary Biology." *Biology and Philosophy* 37 (3): 1–21. <https://doi.org/10.1007/S10539-022-09848-Z/FIGURES/1>.
- Lennox, James G. 1991. "Darwinian Thought Experiments: A Function for Just-so Stories." In *Thought Experiments in Science and Philosophy*, edited by T. Horowitz and G. Massey, 223–46. Savage, MD: Rowman and Littlefield.
- Machamer, Peter, Lindley Darden, and Carl F Craver. 2000. "Thinking About Mechanisms." *Philosophy of Science* 67:1–25.
- Mink, Louis O. 1966. "The Autonomy of Historical Understanding." *History and Theory* 5 (1): 24–47. <https://doi.org/10.2307/2504434>.
- Morgan, Mary S. 2017. "Narrative Ordering and Explanation." *Studies in History and Philosophy of Science Part A* 62 (April):86–97.
<https://doi.org/10.1016/J.SHPSA.2017.03.006>.
- . 2022. "Narrative: A General-Purpose Technology for Science." In *Narrative Science: Reasoning, Representing and Knowing since 1800*, edited by Dominic J. Berry, Kim M. Hajek, and Mary S. Morgan, 3–30. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781009004329.002>.
- Olmos, Paula. 2022. "Just-so What?" In *Narrative Science: Reasoning, Representing, and Knowing since 1800*, 424–44. Cambridge University Press.
- Reydon, Thomas A.C. 2012. "How-Possibly Explanations as Genuine Explanations and Helpful Heuristics: A Comment on Forber." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 43 (1): 302–10.
<https://doi.org/10.1016/j.shpsc.2011.10.015>.
- Richards, Robert J. 1992. "The Structure of Narrative Explanation in History and Biology." In *History and Evolution*, edited by Matthew H. Nitecki and Doris V. Nitecki, 19–55. SUNY Press.
- Ross, Lauren N. 2019. "Causal Control: A Rationale for Causal Selection." In *Philosophical Perspectives on Causal Reasoning in Biology*. Vol. XXI. Minnesota Studies in Philosophy of Science.
- Roth, Paul A. 2017. "Essentially Narrative Explanations." *Studies in History and Philosophy of Science Part A* 62 (April):42–50.
<https://doi.org/10.1016/J.SHPSA.2017.03.008>.

- Roth, Paul A. 2020. *The Philosophical Structure of Historical Explanation*. Northwestern University Press.
- Stanford, P. Kyle. 2017. "The Difference Between Ice Cream and Nazis: Moral Externalization and the Evolution of Human Cooperation." *Behavioral and Brain Sciences*, no. May, 1–57. <https://doi.org/10.1017/S0140525X17001911>.
- Sterelny, Kim. 2012. *The Evolved Apprentice: How Evolution Made Humans Unique*. Jean Nicod Lectures. Cambridge, MA: MIT Press.
- . 2016. "Contingency and History." *Philosophy of Science* 83 (4): 521–39. <https://doi.org/10.1086/687260>.
- . 2021. *The Pleistocene Social Contract: Culture and Cooperation in Human Evolution*. Oxford University Press.
- Sterelny, Kim, and Ronald J. Planer. 2021. *From Signal to Symbol: The Evolution of Language*. MIT Press.
- Swaim, Daniel G. 2019. "The Roles of Possibility and Mechanism in Narrative Explanation." *Philosophy of Science* 86 (5): 858–68. <https://doi.org/10.1086/705528>.
- Tomasello, Michael. 2014. *A Natural History of Human Thinking*. Cambridge, Massachusetts: Harvard University Press.
- . 2016. *A Natural History of Human Morality*. Harvard University Press.
- . 2019. *Becoming Human: A Theory of Ontogeny*. Cambridge, MA: Belknap Press.
- Turner, Derek. 2005. "Local Underdetermination in Historical Science." *Philosophy of Science* 72 (1): 209–30. <https://doi.org/10.1086/426851>.
- . 2009. "Beyond Detective Work: Empirical Testing in Paleontology." In *The Paleobiological Revolution: Essays on the Growth of Modern Paleontology*, edited by David Sepkoski and Michael Ruse, 0. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226748597.003.0011>.
- Turner, Derek, and Ahmed AboHamad. 2023. "Narrative Explanation and Non-Epistemic Value." *Journal of the Philosophy of History* 17 (1): 53–76. <https://doi.org/10.1163/18722636-12341489>.
- Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*. New York, New York: Oxford University Press.
- . 2010. "Causation in Biology: Stability, Specificity, and the Choice of Levels of Explanation." *Biology and Philosophy* 25 (3): 287–318. <https://doi.org/10.1007/s10539-010-9200-z>.