Explaining Transitions in Human Behavioural Evolution: An Interventionist Perspective

Journal:	*Evolutionary Human Sciences*
Manuscript ID	EHS-2023-0082.R1
Manuscript Type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Planer, Ronald; University of Wollongong Pain, Ross; University of Bristol, Philosophy
Keywords:	interventionism, transitions in human evolution, cultural complexity, causation, single-factor explanations
Abstract:	Transitions in human evolution (e.g., the appearance of a novel technological industry) are typically complex events involving change at both spatial and temporal scales. As such, we expect them to have multiple causes. Yet it is commonplace for theorists to prioritise a single causal factor (e.g., cognitive change) in explaining these events. One rationale for this is pragmatic: theorists are specialised in a particular area—say, lithics or cognitive psychology—and so focus on one particular cause, holding all others equal. But could single-factor explanations ever be justified on objective grounds? In this article, we explore this latter idea using a highly influential theory of causation from the philosophy of science literature; namely, interventionism. This theory defines causation in a minimal way, and then draws a range of distinctions among causes, producing a range of different causal concepts. We outline some of these distinctions and show how they can be used to articulate when privileging one cause among many is objectively justified—and, by extension, when it is not. We suggest the interventionist theory of causation is thus a useful tool for theorists developing causal explanations for human behavioural evolution.

Explaining Transitions in Human Behavioural Evolution: An Interventionist Perspective

Abstract: Transitions in human evolution (e.g., the appearance of a novel technological industry) are typically complex events involving change at both spatial and temporal scales. As such, we expect them to have multiple causes. Yet it is commonplace for theorists to prioritise a single causal factor (e.g., cognitive change) in explaining these events. One rationale for this is pragmatic: theorists are specialised in a particular area—say, lithics or cognitive psychology—and so focus on one particular cause, holding all others equal. But could single-factor explanations ever be justified on objective grounds? In this article, we explore this latter idea using a highly influential theory of causation from the philosophy of science literature; namely, *interventionism*. This theory defines causation in a minimal way, and then draws a range of distinctions among causes, producing a range of different causal concepts. We outline some of these distinctions and show how they can be used to articulate when privileging one cause among many is objectively justified—and, by extension, when it is not. We suggest the interventionist theory of causation is thus a useful tool for theorists developing causal explanations for human behavioural evolution.

Keywords: interventionism; single-factor explanations; actual difference making causation; specific causation; cultural complexity; behavioural modernity

1. Introduction

Theorists working on human behavioural evolution often focus on a single causal factor in attempting to explain a target phenomenon (e.g., the origins of the Oldowan). However, these same theorists would typically accept that events as complex as a transition in human behavioural evolution are the product of a broad range of causes. What is going on here? One rationale for this explanatory strategy is pragmatic. A cognitive archaeologist, for instance, may focus on some presumed cognitive cause of a transition (e.g. Coolidge & Wynn, 2018; Mithen, 1996), whereas statistically inclined modellers might focus on population-level dynamics (e.g. Powell et al., 2009; Premo & Kuhn, 2010). Again, no doubt these theorists think that other causes are at play. But they focus on and emphasise one causal factor, or one *type* of causal factor, because that is where their expertise lies. This illustrates the “all-else-equal” strategy in causal inference: the role of one factor is examined closely, and others are held fixed and treated as background conditions. Pragmatic considerations provide a perfectly good rationale for pursuing single-factor explanations, and reflect the practical reality of understanding complex events in the past using a variety of disciplinary perspectives.

But might privileging a single factor in an explanation of something as complex as a transition in human behavioural evolution ever make sense on more than pragmatic grounds? Might there sometimes be an *objective* justification? Recent work in the philosophy of causation on the *interventionist framework* (Waters, 2007; Woodward, 2003) offers a way of exploring this idea. In this literature, a key distinction is made between causes that actually vary and causes that do not in some spatial and/or temporal population of interest. This causal distinction can then be used to carve out still additional causal notions. The main idea is that there can be and often are objective differences among the types of causes at play in a given case. In what follows, we develop this idea in a way appropriate to human behavioural evolution and similar fields, and see what insights follow.

To forestall misunderstanding: we do not think there is anything wrong with being guided by pragmatic reasons in providing a single-factor explanation of some phenomenon. As such, we do not intend to suggest that there is an epistemological or methodological problem for any particular debate in the literature on human behavioural evolution. Our goals are more modest. We aim to outline a basic version of interventionism, and make some initial suggestions regarding its potential application to some much discussed empirical cases. We think the framework developed offers researchers a novel alternative strategy for assessing (and perhaps pursuing) single-factor explanations, but we leave the more fine-grained details of applying the framework to specific empirical cases to those better-versed in the empirical literature.

Returning to the causal framework, interventionists conceive of causation *simpliciter* as follows: two variables, X and Y, stand in a causal relation to one another just in case there are background circumstances in which it is possible to bring about a change in the value of Y by intervening on the value of X. (This is for deterministic cases; in probabilistic cases, the basic idea is that an intervention on X causes a change in the probability distribution over Y.). The notion of an "intervention" here is a specialist one. An *intervention* on X with respect to Y is a manipulation of the value of X such that, if the value of Y is changed, the change in Y only occurs through the change in X. The change in Y cannot occur via some alternative causal pathway that does not include X, for instance, a confounding variable (for a fuller discussion, see Woodward 2003: p. 98).

So, to give a very simple example, your diet is a cause of your weight because there are background conditions in which it is possible to change how much you weigh by changing (or *intervening* on) your diet. This account of causation is extremely minimal, and as such recognizes many relations among variables as cases of causation, some of which might seem counterintuitive at first. Basically*: can you wiggle one variable by wiggling another variable in certain conditions?* However, defenders of the interventionist theory see this as a feature rather than a bug, and we agree. What it provides is an inclusive starting point from which more complex causal notions can be "built up," with theoretically useful relations holding among these notions. For interventionists, this feature is key to advancing a range of issues relating to causal explanation in the sciences.

Below, we begin by setting out a group of causal concepts that help us understand causal relations spread out over space and/or time—concepts that are thus useful in the case of human evolution. Developed by Waters (2007) to deal with some issues in the philosophy of biology, these are the concepts of *the actual difference making cause*, *an actual difference making cause*, and *a potential difference making cause*. We then combine the notion of *specificity* with these concepts, as Waters (2007) also does. The notion of a specific cause has been discussed for some time in philosophy, though often under other descriptions (e.g., "influence"), and often as an account of causation *proper* (e.g., Lewis 2004) as opposed to a *type* of causation. In more recent years, Woodward has neatly formulated the notion of specificity in interventionist terms (see, especially, Woodward 2010). In such terms: a specific causal relationship is one in which *fine-grained* changes can be made to the effect variable by making *fine-grained changes* to the cause variable. After we unpack these causal concepts, we apply them to some key debates in the human evolution literature. In particular, we seek to clearly specify when elevating one causal factor over others in an explanation is justified on objective grounds, and when it is not. We take as our main working example the origins of behavioural modernity. A variety of competing single-factor explanations can be found in the literature on behavioural modernity, including *biological explanations*, *social explanations*, and *environmental explanations*. Our aim is to show both *that* and *how* the interventionist framework bears fruit when applied to this transition. In so doing, we hope to persuade theorists that these causal concepts can do genuine explanatory work in the domain of human evolution.

The discussion will proceed as follows. Section 2 develops the concept of actual difference making causation. Section 3 makes a first pass at applying this thinking to causal explanation in human evolution. Section 4 outlines the notion of causal specificity. In section 5 we bring all of these concepts together, and show how they shed new light on disputes between theorists prioritising biological, social or environmental accounts. We finish by discussing our analysis, and draw some lessons for future inquiry.

2. Actual Difference Making Causation

Imagine you're watching someone light a fire. They take a match out of a matchbox and strike it under a pile of wood. The wood ignites. At the start of this process, there was no fire; now there is. Finding yourself in a philosophical mood, you wonder *what caused the fire to light*?

The striking of the match, you think. And of course, you're right. But are matters really so simple? Most obviously, you recall from chemistry class that fire needs oxygen to burn. Hence, you conclude that, were oxygen not present, there would be no fire. Nevertheless, you can't shake the feeling that the striking of the match, rather than the presence of oxygen, is somehow more important. Were someone to ask you, "What caused the fire to light?" and you responded "the presence of oxygen", this would be considered odd. If, however, you responded "the striking of the match", this would be accepted. But why? What exactly is the difference between these two answers? Perhaps this reflects nothing more than an understandable bias: episodes of match striking just grab our attention in a way that the presence of oxygen doesn't. After all, objectively speaking, both the striking of the match and the presence of oxygen are on par with one another, aren't they?

The distinction between actual vs. potential difference making offers a simple but insightful analysis of such cases. To see how the analysis works, let's introduce some variables. Let Strike be a variable having values *yes* and *no* (we use capitalized words for variables and italicized words for values of variables). Strike being set to *yes* corresponds to the state of affairs in which the match is struck; Strike being set to *no*, corresponds to the match *not* being struck. Next, let Fire be a variable, again with values *yes* and *no*, where Fire being set to *yes* corresponds to a state of affairs in which there's fire, and *no* to an absence of fire. The two variables stand in a causal relation to one another, as there are background conditions—indeed, plenty of them—in which we can manipulate the value of Fire (from *no* to *yes*) by manipulating the value of Strike (from *no* to *yes*):

Strike = *yes*, *no* → Fire = *yes, no*.

So far, so good. But now consider a third variable, Oxygen, which also takes values *yes* and *no*. The same reasoning that justifies our treating Strike and Fire as causally related applies here, i.e., there are background conditions (including the ones we imagine to actually hold in our example) in which we can manipulate the value of Fire by manipulating the value of Oxygen:

Oxygen = *yes*, *no* → Fire = *yes*, no.

It is in this sense, then, that it's right to think that both match striking and the presence of oxygen are causes of the fire's lighting. However, while the two variables are on par in this respect, they differ crucially in another. Specifically, while both Strike and Oxygen are causes of Fire, only *one* of these variables *actually varied* in the leadup to the fire's lighting. More precisely, let t signify the period just before the fire was lit; at t, the pile of wood sits unlit in the fireplace. And now let t^* signify the period just as the fire lights. There is thus variation in the value of Fire between t and t^*:

Fire = *no* at t, Fire = *yes* at t^*.

Now, and this is the crucial point: if we look to our two cause variables, Strike and Oxygen, we see that it was only Strike that varied over the relevant timescale, that is, between t and t^*, whereas oxygen was present throughout:

Strike = *no* at t, Strike = *yes* at t^*.

Oxygen = *yes* at t, Oxygen = *yes* at t^*.

So, while both Strike and Oxygen are causes of Fire, it was variation in Strike, and not in Oxygen, that explains the variation in Fire between t and t^*.

Following Waters (2007), we shall say that Strike was *the actual difference making cause* of Fire, while Oxygen was merely a *potential difference making cause*. To be a potential difference maker with respect to some effect variable in a given context, it is enough to simply be a cause of that variable (in the interventionist sense) in that context. On any natural way of filling out our example with more details (as we do later), there would be many other potential difference making causes of Fire in this case. The variable Dry, for instance, specifying whether the matchstick is dry (*yes*) or wet (*no*), is an obvious example. Just as if oxygen had not been present, the fire would not have lit, so too if the matchstick had been wet, the fire would not have lit. Or another: the variable Matchbox, specifying whether there is an appropriate surface for the match to be struck on (*yes*), or not (*no*).

But often it will be the case that multiple causes actually vary. Then, there is no one causal factor that is *the* actual difference maker; instead, what we have is a group of actual difference makers, each of which is *an* actual difference maker. To see this, consider another case. You're watching someone attempt to light a burner on a gas stove. They turn the knob under the burner and strike a match just beneath the burner. The burner ignites. There is variation over time in the effect variable Fire: at t, there's no flame; at t^*, there is. But now, there is variation in not one but two causes of Fire: Knob (*on*, *off*) and Strike (*yes*, *no*).

(Knob = *off* at t, *on* at t^*; Strike = *no* at t, yes at t^*) → Fire = *no* at t, *yes* at t^*.

In this case, both Knob and Strike are each *an actual difference making cause* of Fire (while Oxygen remains a potential difference maker).

We think these simple and intuitive distinctions can help clarify theorising about transitions in human behavioural evolution. Reflecting on the questions posed at the start of this article (and foreshadowing the discussion to come): note that there are clearly aspects to this way of understanding causation that are relative to our interests as explainers. Some of these aspects are simply inherited from the interventionist theory of causation. Even if the world came neatly pre-packaged into variables of the sort that appear in interventionist models, we would still face the task of selecting some variables rather than others in attempting to understand the causal structure of a target system. This includes the selection of a particular effect variable (or variables), the change(s) in which we seek to understand. What we wish to explain in the first place depends on our interests. In other words, pragmatic rationales are always at play. But once all these choices are made, it is objective features of the target system that determine actual vs. potential difference making causation. Returning to our first example: the match was unstruck at t, and then struck at t^*, while the presence of oxygen remained constant throughout. Identifying match-striking as the cause of the fire thus picks out an *objective fact* about change in the structure of the world over time, and likewise offers an objective ground for an explanation prioritising this cause.

Let's now begin to look at how these ideas can throw light on causal-explanatory issues in human evolution.

3 Explaining Transitions in Human Behavioural Evolution: Part A

Theorists have developed a range of explanations for transitions in human behavioural evolution. Paradigm examples of such transitions include various increases in cultural complexity: the appearance of a novel lithic industry (e.g., the Acheulean, Levallois), directional change within such an industry (e.g., change from the Early to the Developed Oldowan), the appearance of so-called symbolic behaviours (e.g., adorning the body with jewellery, pigments, cave painting, etc.), and so on. Here we sort commonly cited causal factors into three categories: biological, social, and environmental.

3.1 Biological factors

We can abstractly represent the causal relations posited by this category of explanations as taking the form:

Biological Factor B = x, y → Cultural Factor C = x, y.

This is a kind of *schema* that biological explanations fill in with more concrete variables (and values).

Suppose that the cultural factor in question is, e.g., the appearance of the Acheulean handaxe in the archaeological record. Our effect variable is: Handaxe = *present, absent*. In terms of biological causes of this effect, a number have been offered. What is perhaps most impressive about handaxes, relative to earlier Oldowan tools, is the complexity of their design and hence the increased cognitive demands handaxes placed on their makers. Hypothesized causes have included enhanced working memory capacities (Coolidge & Wynn, 2018) and/or enhanced capacities for hierarchical cognition (Stout et al., 2021; Stout & Chaminade, 2012). In addition, some link the handaxe industry to the origins of novel social learning abilities, such as imitation learning (Arbib, 2011; Paddayya, 2004).

Biological factors are typically envisaged by theorists as intrinsic traits of hominin minds (and/or bodies). By "intrinsic," we mean these traits do not depend on specific environmental conditions for their acquisition or development; they are *robustly developing* traits (Northcott and Pinccinini 2018; see, also, Ariew 1999). These traits are understood as largely genetically specified or "coded" (though it would probably be better to think in terms of genetic canalization (Waddington 1945) here); hence, their appearance is understood as the result of a genetic mutation, while their establishment at the population level is explained in terms of natural selection operating on genes. Some biological factors, including some cognitive ones (see, e.g., Heyes 2018 on "cognitive gadgets"), instead owe to mechanisms of adaptive plasticity. However, here we delimit our focus to the more common understanding of "biological factor" in the literature (i.e., a strongly genetically canalized or channelled trait).

3.2 Social factors

Similarly, we can schematically represent social causes like this:

Social Factor S = *x*, *y* → Cultural Factor C = *x*, *y*.

In this category are causes relating social dynamics within and/or between hominin social groups to cultural changes. Such factors are typically envisaged as capable of undergoing change independently of changes to hominins' intrinsic biological (including cognitive) traits: holding these traits fixed at a time, hominin social networks can expand, contract, change in their internal composition, etc. (though these changes may well produce downstream biological change). The highly influential demographic models that link hominin cultural complexity to features of social learning networks (in particular, their effective size) squarely fit in this category (e.g. Powell et al., 2009; Premo & Kuhn, 2010). So do models that emphasize not just the impact of population size, but also the unique ways in which factors like migration and meta-band structure influence cultural innovation and the spread of innovations (see, especially, Sterelny 2021a, 2021b). A different though related line of thought that goes here connects the establishment of cooperative breeding to increases in cultural complexity. Cooperative breeding increases the social complexity of hominin lives in a variety of ways, including in some that directly bear upon social learning. It has been plausibly argued, for example, that cooperative breeding provides learners with a larger pool of tolerant, in-group models to learn from (as opposed to, say, just one's mother and/or siblings) (Burkart et al., 2009; Burkart & van Schaik, 2016; Hrdy, 2011).

3.3 Environmental factors

Finally, we can think of environmental explanations as taking the form:

Environmental Factor E = x, y → Cultural Factor C = x, y.

For example, climatic instability both at and over time has been used to explain transitions in human behavioural evolution, including cultural complexity (e.g. Richerson & Boyd, 2013; Shultziner et al., 2010). Another highly influential idea has been that the *risk* associated with particular environments promotes increases in cultural complexity (Collard et al., 2005, 2013, 2016; Torrence, 1983, 1989, 2001). Environments differ with respect to the risks (e.g., risks of resource failure) that they pose, sometimes sharply. According to this hypothesis, high risk of resource failure selects for a more complex and diverse tool kit. This is because, in such environments, the costs of a technological misfire tend to be dire. As such, foragers are expected to develop tools that are more specialized and (hence) more reliable; tools that better mitigate risk. Specialized tools, in turn, tend to be more internally complex (i.e., have a greater number of functional parts) and more complex to manufacture.

3.4 Actual vs. potential difference making causes of transitions in cultural complexity

How do these various explanations look from the interventionist perspective outlined above?

We think factors from all three of the above categories—the biological, the social, and the environmental—are going to be causally relevant to understanding major transitions in human cultural complexity. But this may be so only in a minimal sense. Namely: there will be instances of each type of factor that indeed serve as *causes* of the effect variable of interest; that is, as potential difference makers. To see this, let's take a step back. Suppose, again, that it is the appearance of the Acheulean handaxe (or the Levallois flaking technique, or some other impressive tool form) that we wish to explain. The rationale for thinking that biological factors feature among the causes seems to run as follows: (i) tools place various task demands on makers' intrinsic cognitive capacities (Intrinsic Cognition); and (ii) agents' intrinsic cognitive capacities depend on agents' biological makeup (for example, gross facts about their brain size and/or organization). Thus, in general, it is possible to intervene on the set of artifacts an agent can reliably make by intervening on certain properties of their brains—say, those affecting their working memory abilities—and we have every reason to believe this was true in the circumstances that actually accompanied the appearance of the handaxe in the archaeological record; in particular, the appearance of *Homo erectus* (de la Torre, 2016). In this sense, it can be truly said that it takes a mind of a particular intrinsic grade to manufacture an Acheulean handaxe. Intrinsic Cognition is among the *bona fide* causes of Handaxe.

The same can be said for both social and environmental factors, in our view, though the reasons for recognizing causal links between these sorts of factors and changes in cultural complexity are different. More specifically, here, the grounds for positing causal relations are provided by a mix of theoretical, modelling, and empirical (e.g., ethnographic) evidence. This evidence shows there is indeed a wide range of empirically realistic circumstances in which, by intervening so as to change (e.g.) population density or (e.g.) the risks of resource failure, we can bring about change in a population's material culture. All that is required for population density or resource risk to have been a genuine cause of some transition in cultural complexity is for the actual background circumstances that held at the time to have supported the relevant counterfactual relations (i.e., that the effect variable Cultural Factor C would have differed in its value had Population Density and/or Risk varied in its value).

That said, it will by now be clear that being a cause *simpliciter*, or a potential difference maker, is one thing, while being *the cause that actually made the difference* is another. We propose that a hypothesis holding that, for example, it was Intrinsic Cognition, and not Population Size or Risk, that drove a particular cultural transition can be understood as the claim that Intrinsic Cognition was the actual difference making cause of the transition, whereas these other variables merely served as potential difference making causes. In this way, biological, social, and environmental explanations can indeed *compete* with one another, as the remarks of theorists often suggest (see Section 5.2 below), even if all serve as genuine causes of some transition.

Accordingly, when there is variation in just a single cause variable (or better: in just a single cause among a family of causes under consideration) with respect to some effect variable of interest, this will in general provide an objective reason for singling that cause out among the wider group of variables all of which are causes of the effect. In principle this need not be any different than explaining the fact that the fire lit by citing the striking of the match. In practice, of course, there are bound to be serious methodological issues in coming to know this, but perhaps not insurmountable ones. In any case, we think having these distinctions between causal properties is a useful idealisation, which real-life cases can approximate to a greater or lesser extent. In particular, it provides an objective rationale for favouring single-factor accounts in at least this sense: once theorists agree on the effect they wish to explain (which will include specifying the temporal and/or spatial stages across which there is actual variation), and on the broader set of variables that are to be searched among for causes of this effect, it is objective facts about the structure of the world that determine which factor it is that actually made the difference to the effect in question. We suspect a failure to agree on these conditions might lead theorists to talk past each other. There need be no in principle disagreement, for example, between theorists who think that social factors are the actual difference making cause of some transition and theorists who focus on biological factors for pragmatic reasons.

And yet: when dealing with a phenomenon as complex as a transition in human behavioural evolution, it will generally be the case that there is more than one cause variable that actually varies over the relevant temporal and/or spatial scale. For example, both climatic instability and population density might actually vary. Moreover, two or more cause variables with respect to the effect variable may *themselves* causally interact over time; climactic change might causally interact with population density (say). In turn, a change in demographic conditions might then lead to causally relevant variation over time in biological factors through natural selection operating on genotypes. And this becomes increasingly likely as time depth increases. The possibility of interactions among such causal factors at and over time no doubt poses thorny epistemological/methodological challenges (as a referee emphasized to us) for our attempts to reconstruct the causal histories of transitions in human behavioral evolution (though, again, perhaps not insurmountable challenges). This is a topic for another day, however. For now, we limit ourselves to the following issue: the inherent causal complexity of such cases would seem to erode any objective ground for privileging one of these causes over the others. To the extent that such an explanatory strategy *is* justified in such a case, it is natural to think it must be on purely pragmatic grounds. But perhaps this is too fast. Perhaps there are other, objective, reasons why it might make sense to prioritise one cause over others, even where each is an actual difference maker of the effect of interest. We turn to this issue now.

4. Specific Actual Difference Making Causation

We now introduce another causal notion, namely, that of a *specific cause*. The influence or power a cause variable has over an effect variable can be more or less *specific*. The idea of specificity is often expressed in terms of "fine-tuning." Specific (or highly specific) causes are ones whose value you can fine-tune, and in so doing, fine-tune the value of the effect variable. In contrast, non-specific causes operate in a switch-like fashion: you can change the value of the effect variable by intervening on the value of the cause, but you can't modulate the value of the effect variable in a fine-grained way by modulating the value of the cause variable.

A simple adaptation of one of our above examples can be used to illustrate specificity. Recall our example involving the burner on the gas stove. Let us now tweak the example. As we originally described this case, the variable Knob had just two values: *on* and *off*. Let's now enrich the value space of this variable as follows: the knob has an off setting corresponding to 0° of rotation; an ultra-low setting corresponding to 45° rotation, a low setting corresponding to 90° rotation, and so on (see Figure 1). The value space of the variable thus looks like this:

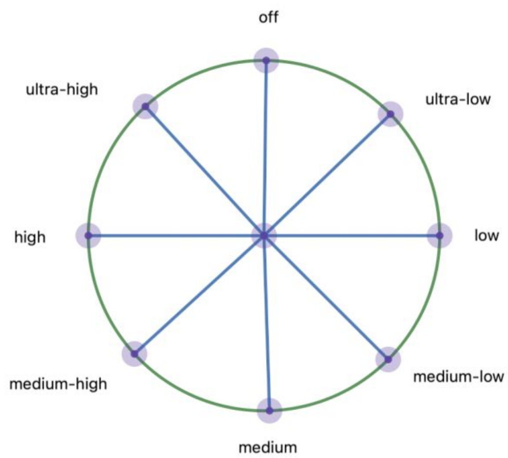
Knob = *0°, 45°, 90°*, ..., *315°*.

Similarly, let us also enrich our description of the burner's state. The flame can be in a variety of states: *absent*, *ultra-low*, *low*, *medium-low*, and so on:

Fire = *absent*, *ultra-low*, *low*, …, *ultra-high*.

Here, we say that there is a *specific causal relation* between Knob and Fire (Woodward, 2003, 2010). For not only is it possible to change the value of Fire by intervening on Knob; it is possible to *fine-tune* the value of Fire by *fine-tuning* the value of Knob.

(a)



(b)

Knob	Fire
0°	*absent*
45°,	*ultra-low*
90°	*low*
135°	*medium-low*
180°	*medium*
225°	*medium-high*
270°	*high*
315°	*ultra-high*

Figure 1. A geometric (a) and numeric (b) specification of value correspondences for Knob and Fire.

This, then, is the essence of causal specificity from an interventionist perspective. Here, however, we are interested in a particular variant on this notion: that of a *specific actual difference making cause* (Waters 2007). To be a specific actual difference maker, it is not enough to be a specific cause *and* an actual difference maker. Put differently, we might say that a cause can be a specific cause of some effect (in the abstract) without exercising specific influence over the effect in some actual situation. This can occur in two ways. First, a cause may be a specific cause of some effect in some background conditions, but not in those that actually obtain in the case under consideration. (A cause can be a specific cause relative to one set of background conditions but not another.) This is compatible with the cause actually varying in a fine-grained way yet nonetheless failing to serve as a specific cause in the case at hand. The conditions that are conducive to the cause functioning in a specific capacity are not actually in place. Second, the background conditions that hold in a given case may be such that, *were* the cause to vary in a specific way, the effect variable would likewise vary in a specific way. And yet, as a matter of actual fact, the cause does not vary in a specific way. It might vary, but only between two values over the relevant temporal/spatial frame (hence, not in a fine-grained way). Or it might not vary at all. To be a specific actual difference maker it's necessary that the effect variable *actually* vary in a (more or less) specific way, that the cause variable likewise *actually* vary in a (more or less) specific way, and that the specific variation in the effect be at least (partially) counterfactually dependent on the variation in the cause.

So, for example, suppose we were to observe the variable Fire to vary over time as follows:

Fire = *absent* at t;

Fire = *medium-low* at $t+n$;

Fire = *high* at $t+n+m$ (for n, $m > 0$).

And now suppose that it's this variation in Fire over time that we wish to explain. Consider, then, the variables: Oxygen, Strike, and Knob. Oxygen is a cause of Fire, but (as before) it is only a potential difference making cause; it does not actually vary over the appropriate timescale. In contrast, not only are Strike and Knob also causes of Fire; both *actually vary* over the appropriate timescale, and so both count as actual difference making causes of variation in Fire. However, only one of these actual difference makers, namely Knob, is a *specific* cause of variation in Fire. It is the only *specific actual difference maker*. The striking of the match, in contrast, acts like a binary switch with respect to Fire. Yes, by intervening on Strike, one can change the value of Fire. But what one cannot do is modulate the value of Fire in a fine-grained way.

The upshot is this: the distinction between specific and non-specific causation provides another reason why it can be legitimate to give special status to one cause among many in an explanation, even when there are multiple actual different makers at play. And like the distinction between actual vs. potential difference making, a variable's status as a specific actual difference maker is based on objective features of the target system. While the *desire to know* the specific cause, or the specific actual difference maker, of some effect is interest-relative, the property of being *the* or *a* specific actual difference maker is not. Consequently, there can be an objective rationale for a single-factor explanation, *even where* multiple causes all serve as actual difference makers.

5. Explaining Transitions in Human Behavioural Evolution: Part B

How can the concept of specific actual difference making contribute to debates in human behavioural evolution? To begin, we note that the social and environmental factors commonly cited to explain cultural complexity—for example, population density and risk—are clearly paradigm cases of specific causes in the above sense. (i) Each is conceived of as a many-valued cause variable, and (ii) background conditions (which are often only implicitly specified) exist in which the cultural complexity of a group (as measured, for example, by the number of tools they possess, or the complexity of individual tools) can be turned up or down by turning up or down the value of these cause variables.

The situation with respect to causal specificity in the biological case is more complex. Some biological causes of cultural complexity are standardly conceived of in non-specific terms. A clear example is the possession of shared intentionality (e.g., Tomasello et al. 2005). A more complicated example is so-called *know-how copying*; a form of social learning encompassing, but not limited to, imitation learning (e.g., Bandini et al., 2020; Bandini & Tennie, 2018; Tennie, 2023). Know-how copying is generally treated as a capacity agents either have or don't have, with its presence or absence being taken to explain the type of culture observed on the part of some group. Debate continues, for example, as to whether any other great ape has the capacity to genuinely copy each other's manual behaviours (for a recent overview, see Whiten (2022), and responses therein). Yet at the same time, one might think it natural to carve up know-how copying into more and less error prone forms (and, in any case, it's easy to make sense of an agent being more or less *disposed* to copy). Other biological causes are similarly or even more open to interpretation. For example, consider the notion of "cognitive fluidity" (Mithen 1996). The cognitively fluid mind is one in which all the mind's (previously informationally isolated) "modules" can talk to one another. A fluid mind can seamlessly weave together information from its "naïve biology" and "naïve physics" modules, so as to create, for example, composite tools featuring both organic (wood, bones) and inorganic (stone) materials. While it is very natural indeed to imagine fluidity coming in degrees, that is not how the idea has generally been developed in the literature. Rather, Mithen and others have treated fluidity as an all-or-nothing intrinsic cognitive trait explained by the evolution of complex syntactic forms of language (itself often conceived of as an all-or-nothing trait, arising due to a sudden genetic mutation) (Berwick & Chomsky, 2016, 2019; Tattersall, 2016). Other cases of biological factors that are regularly understood by theorists as—though they would not use this language—specific causes of cultural complexity. Paradigm examples include working memory capacities (e.g. Wynn & Coolidge, 2004), hierarchical cognitive capacities (e.g. Stout et al., 2021; Stout & Chaminade, 2012), and orders of intentionality (e.g. Cole, 2016, 2019).

5.1 The origins of behavioural modernity

To illustrate these causal concepts in a more concrete way, we now zoom in on a particular transition in human cultural complexity, namely, the origins of behavioural modernity. By this, we have in mind the suite of behavioural, and specifically, cultural traits that are either unique to modern *sapiens*, or which are at least uniquely highly developed or prevalent in *sapiens*, compared to our Neanderthal and Denisovan cousins. While there remains significant debate among archaeologists in this area (see, e.g., Nowell (2010) and references therein on the important nuances in debates over the technological and symbolic differences between *sapiens* and Neanderthals in relation to behavioral modernity), this controversy mostly takes place below the level of what is relevant for our purposes. In our view, there is enough agreement regarding the existence both of an interesting set of technological and symbolic differences between *sapiens* and Neanderthals, as well as the timeline of the establishment of these differences, for the case study to be a useful and illuminating one.

As is well known, for a long-time orthodoxy held that the behaviourally modern package emerged suddenly in Europe around 50 kya in what was often referred to as an "explosion" or "revolution." The thought was that modern humans living in this region had rapidly evolved forms of culture on par with those of ethnographically known foragers. These *sapiens* were equipped with new and highly sophisticated technological forms—blades, composite tools, true projectile weapons—as well as elaborate symbolic forms, which were presumably used to navigate much more complex social worlds. In contrast, Neanderthals, who had lived in the same region for hundreds of thousands of years, had evolved few or none of these signature signs of modernity.

The (supposed) sudden onset of behavioural modernity in Europe was highly salient from the perspective of this early consensus, which fuelled belief in a biological cause: a chance genetic mutation had arose and rapidly spread to fixation among these *sapiens* (Coolidge & Wynn, 2018; Mellars, 2005; Mellars & Stringer, 1989). But in addition, archaeologists saw a tight connection between the complexity and sophistication of many of these novel cultural forms and intrinsic cognition. The manufacture of elaborate composite tools and cave paintings depicting supernatural entities required a new kind of mind. It was an intrinsic cognitive change that provided the "spark" that ignited the Upper Palaeolithic "explosion."

Such a sudden origins scenario for behavioural modernity is all but universally rejected nowadays. In their landmark paper, "The Revolution that Wasn't," Mcbreaty and Brooks (2000) made the case that many of the elements of the behaviourally modern package—for example, microliths, body pigments, jewellery and art—instead appear in Africa tens of thousands of years earlier. Subsequent archaeological research strongly confirmed their gradualist account. More specifically, it is now widely agreed that many of these same innovations show a patchy temporal distribution: they appear in a region, last for a period, disappear from that region, and then reappear at some later time, presumably having been re-innovated by *sapiens* (Hiscock & O'Connor, 2006). Explaining this new data eventually motivated a second—and in many quarters, still dominant—wave of explanations for the onset of modernity, this time revolving around demographic factors (Boyd, 2017; Henrich, 2015; Muthukrishna & Henrich, 2016; Powell et al., 2009; Richerson & Boyd, 2008, 2013). To the extent that changes in demographic variables could explain changes in cultural complexity, this seemed like a much more plausible explanation of the transition. For it is easy to see how factors like population density could wax and wane over time.

Finally, and more recently, this demographic consensus has been strongly criticized by proponents of an environmental risk explanation (Collard et al., 2005, 2013, 2016; see in particular 2016: p.2). With these archaeologists, the focus is primarily on technological complexity, but there are also views that connect resource strain and other crises to an expanded role for symbolism in *sapiens* groups (e.g., Straus, 2000). This line of thought is supported by general behavioural-ecological conditions, but more importantly, also by a range of empirical surveys examining the complexity of hunter-gatherer tool kits under varying conditions of risk. The key point is this: like population density, risk is something that can vary, not just over space, but also over time due to shifting climactic conditions. Like population density, risk can rise and fall.

5.2 Applying the causal concepts

To bring the above causal tools to bear on this transition, the first thing we need is a clear specification of our effect. This can be Behavioural Modernity. And we'll let this variable take the values *absent, partially present, and fully present*. Obviously, this is a *massive* oversimplification. But this will be all we need to make our central points.

As it's actual variation in this variable that we want to explain, the next thing to do is clearly specify the relevant population. Here, we'll just focus on time. Had behavioural modernity in fact appeared suddenly, as orthodoxy originally maintained, only two temporal stages would have been necessary, *t* (Behavioural Modernity = *absent*) and *t** (Behavioural Modernity = *present*). Now we know that won't do. We propose something like the below pattern of variation over time (Table 1) to stand for the patchy onset of behavioural modernity.

Time	Behavioural Modernity
250 kya	*absent*
200 kya	*partially present*
150 kya	*absent*
100 kya	*partially present*
50 kya	*fully present*

Table 1: A highly simplified depiction of the patchy emergence of behavioural modernity over time.

We now face three questions: (i) what are the causes (i.e., potential difference makers) of this effect; (ii) what are the actual difference making causes of this effect; and (iii) what (if any) are the specific actual difference making causes of this effect? We emphasize that our aim here is to illustrate possibilities, not defend particular answers to these questions. The latter would, among other things, require a much more empirically realistic setup than we are working with here.

5.2.1 Intrinsic Cognition

We agree with those cognitive archaeologists who emphasize the cognitive demands of many of the cultural forms associated with behavioural modernity. The artifacts and symbols symptomatic of modernity *really are* impressive from a cognitive point of view. We doubt, for example, that erectine minds were capable of innovating bow-and-arrow technology. Whether heidelbergensians, with their much larger brains, might have done so is a more difficult question. Perhaps, but perhaps the elaborate forms of symbolism known from caves like Chauvet were beyond their cognitive reach. The important point is that behavioural modernity depended on a sophisticated cognitive platform (e.g., modern or like-modern working memory capacities). If so, then *sapiens*' Intrinsic Cognition was indeed a *bona fide* cause of behavioural modernity.

The real question, in our view, is whether Intrinsic Cognition *actually* varied over the relevant timeframe in a way that might explain the origins of behavioural modernity. In other words, was Intrinsic Cognition an actual difference making cause? For many theorists, the "sudden appearance" of the behaviourally modern package in Europe between 50-40 kya was by far the most compelling argument for a biological (ultimately, genetic) explanation for behavioural modernity. But with this origins scenario superseded by subsequent finds, many theorists now express strong skepticism over such an explanation. For example, Sterelny (2014) writes:

> … the material traces of modernity are much less stable than we would expect, if those traces are the social reflections of a distinctive and genetically canalized set of enhanced cognitive capacities" (p. 67).

He then continues (in a footnote):

> Of course, it would still be possible to suggest that the genetic change was necessary but not sufficient for modernity. But this would rob the explanatory strategy of its interest, both because of the lack of a positive case for the idea, and because attention would shift to identifying the extra factors, presumably to do with social complexity. (*ibid.*)

We find two things noteworthy about these remarks. First, we note that talk of "necessary conditions" in this context is ambiguous. It is ambiguous between a factor's merely being a cause *simpliciter*, that is, a potential difference making cause of behavioural modernity and its being an actual difference making cause. These represent two different objective scenarios. Second, we think Sterelny's claim that "attention would then shift to identifying the extra factors" is best understood as an implicit request for the specific actual difference making cause (or causes) of behavioural modernity.

We view it as an open question whether Intrinsic Cognition was an actual difference making cause of behavioural modernity. But even if it was, it's clear that it could not have been *the* or even *a* specific actual difference making cause. This is true even if the specific form of Intrinsic Cognition that is envisaged to have played a role in the transition is itself a specific cause of cultural complexity (e.g., working memory capacity). This is for the simple reason that no Intrinsic Cognitive factor can be expected to appear, then disappear, then reappear again (etc.) over the 200 ky timescale over which behavioural modernity establishes. Such a scenario would be completely outlandish from a biological perspective (though this would be different if the cognitive capacity in question were instead *culturally constructed*, as in the sense of Heyes, 2018).

5.2.2 Population density

Let us now turn to social factors. Our read of the literature is that many theorists agree that population density fluctuated over the last 250 ka, and that such variation, as indicated by the formal models, can explain the gradual and patchy onset of behavioural modernity. What is the positive evidence in support of the hypothesized fluctuations in population density that might drive this change? Here is one example: Cieri et al. (2014) have plausibly connected changes in *sapiens* craniofacial anatomy (what they call "feminization") to increased levels of social tolerance in *sapiens* over the last 200 ka.

The crucial link concerns reduced levels/effects of circulating testosterone in adults. Such increased social tolerance is plausibly understood as an effect (and possibly a cause) of increased levels of population density. But interestingly, in this case, the pattern is not one of population density waxing and waning (as reflected in craniofacial anatomy), but of steadily being on the rise over this time period. (At present, genetic studies paint a complex, changing, and often conflicting portrait in this area: see, e.g., Li and Durbin 2011; Schlebusch et al. 2012; Schiffles and Durbin 2014; Bergström et al., 2021.)

So, whereas in the case of Intrinsic Cognition we are inclined to think the main question is whether Intrinsic Cognition was an actual difference maker or merely a potential difference maker, here, we are inclined to think that the main question is whether Population Density was a specific actual difference maker, or simply an actual difference maker.

5.2.3 Risk

Finally, let's consider risk as a paradigm environmental factor. Beginning around 800 kya, the Earth entered a phase of marked climactic instability characterized by alternating periods of warming and cooling (Figure 2). This pattern reached its peak over the last several hundred thousand years. With such fluctuation in climactic conditions, we would expect the risk of (e.g.) resource failure in a region to likewise fluctuate over time. In Africa, colder temperatures would have led to more arid conditions, leading to a reduction in primary biomass and hence food for foragers.

The idea that a causal link obtains between risk and cultural complexity has received increasing empirical support in recent years. More specifically, a number of surveys focused on hunter-gather groups (as opposed to, say, farming or horticultural societies) have found that risk is a better predictor of tool-kit complexity than population density. These results are nicely summarized in Collard et al. (2016). They go on to conclude:

> That more than two-thirds of the tests of the population size hypothesis that have been carried out to date do not support the hypothesis casts doubt on its use to explain patterns in the archaeological record … Given that not even a majority of studies indicate that population size is the dominant driver of cultural complexity, there are no grounds for invoking population size to explain patterns in the archaeological record. (p. 6)

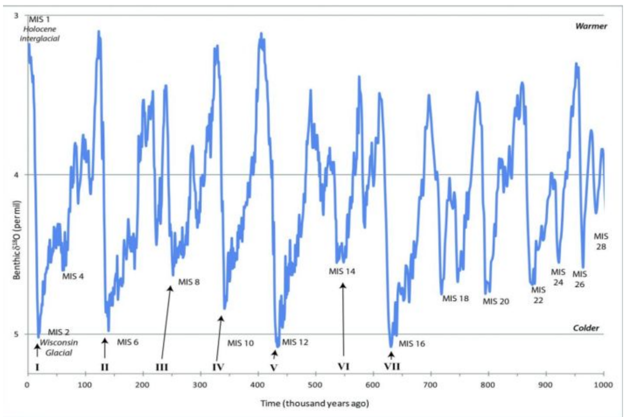


Figure 2: Marine Isotope Stages (last 800 kya). Taken from: Lisiecki, L. E., and M. E. Raymo (2005), A Pliocene-Pleistocene stack of 57 globally distributed benthic d180 records, Paleoceanography, 20, PA 1003, do:10.1029/2004PA001071.

We suspect that what Collard et al. have in mind by "dominant driver of cultural complexity" here is quite close, if not identical to, being the specific actual difference maker of cultural complexity in a given case. We suspect that in claiming that there are "no grounds" for appealing to population density explanations of cultural complexity in the archaeological record, what they mean is these studies provide no reason for thinking population density is the specific actual difference maker of cultural complexity. But at the same time, in claiming that population density is not the "dominant driver", Collard et al. appear to be making room for the idea that social factors might play some role, just not the most important role. If this is correct, it would be useful for all this to be made explicit. The causal concepts and distinctions outlined here are, we think, well suited to such a theoretical task.

Here is one way all of the above types of factors might fit together in an evolutionary scenario, then. Again, we emphasise that our goal here is mainly to illustrate possibilities, rather than to defend this particular scenario.

- Intrinsic Cognition is a potential difference maker of Behavioural Modernity;
- Population Density is an actual difference maker of Behavioural Modernity; and
- Risk is a specific actual difference maker of Behavioural Modernity.

Time	Intrinsic Cognition (= Potential Difference Maker)	Population Density (= Non-Specific Actual Difference Maker)	**Risk (= Specific Actual Difference Maker)**	Behavioural Modernity
250 kya	*present*	*low*	***low***	*absent*
200 kya	*present*	*medium*	***medium***	*partially present*
150 kya	*present*	*medium*	***low***	*absent*
100 kya	*present*	*medium*	***medium***	*partially present*
50 kya	*present*	*high*	***high***	*fully present*

Table 2: The specific actual difference maker (Risk) is shown in bold.

Table (2) depicts this scenario. On this hypothesis, there is no actual variation in Intrinsic Cognition; there is actual variation in Population Density, but not of a specific variety (see the row headed by 150 kya); while there is specific actual variation in Risk.

6. Conclusions and Future Directions

Let's take stock. We have seen how the elevation of one cause among many in an explanation of some phenomenon can be justified by objective facts. The strongest case for a single-factor explanation is when there is just a single actual difference making cause (specific or not). The phenomenon may have many causes, that is, potential difference making causes. However, if and to the extent that there is just a single cause that *actually* varies, and in so doing, *actually* brings about the change in the effect variable that is of interest to us, it can be entirely sensible to single out this cause, or otherwise give it a place of explanatory prominence, among many. The special status of this causal factor is grounded in objective features of the target system: this cause actually varied, whereas the others didn't.

As we move out from simple and straightforward cases, matters grow correspondingly more complex. What justification is there for single-factor explanations when there is more than a single actual difference maker, as will often be the case for phenomenon as complex as transitions in human behavioural evolution? If there is no objective asymmetry between actual causes, then one might think we only have pragmatic reasons. However, if there *is* such an asymmetry—if, for example, only one of the actual causes functions in a specific capacity, whereas the other(s) function in a (more) switch-like fashion—then a single factor explanation can again seem entirely reasonable.

We doubt there is a one-size-fits-all story to tell at this point. In particular, sometimes a non-specific actual difference maker may serve as a crucial background condition for a specific actual difference maker (e.g., Population Density for a specific Risk → Behavioural Modernity causal link, or perhaps Intrinsic Cognition for a specific Population Density → Behavioural Modernity causal link). When this is the case, an explanation that fails to cite the non-specific actual difference maker may be felt to be inadequate or misleading. This is especially true if the specific actual difference maker varied in a specific way *prior* to the appearance of the effect. For example, plenty of Mid-Pleistocene hominin groups encountered high levels of resource risk without innovating modern tools. One might think this is precisely because population density had not yet reached the threshold required to allow the (putatively specific) Risk → Behavioural Modernity causal link to operate.

The central takeaway is this: it is important to be clear on the ontological status of the various (hypothesized) causes in a given transition in human behavioural evolution—potential difference maker, actual difference maker, and specific actual difference maker. This, alone, should go a significant way towards helping clarify certain debates over how to best explain and understand transitions in human behavioural evolution.

There are various directions for future research based on the framework developed here. We contend that the interventionism can help pinpoint exactly what is at issue between competing causal hypotheses in human evolutionary theory. However, the framework we have suggested here is very basic. A lot more work is required to think through how the causal concepts outlined might apply to specific debates in all their detail. In particular, future work would focus on applying the framework in a manner that takes into account the full empirical complexities of a particular debate. For instance, it would be interesting to focus on the debate between proponents of environmental models (Collard et al., 2005, 2013, 2016) and social models (Boyd, 2017; Muthukrishna & Henrich, 2016; Powell et al., 2009) in much more detail. Exactly where does the conflict between them lie, and where they conflict, can we adjudicate between them? Have avenues for fruitful synthesis been overlooked? Of interest here is that the variation in the former models tends to be spatial, whereas the variation in the latter tends to be temporal. The evidence adduced by Collard and colleagues primarily concerns variation in forager kit across spatially disparate forager populations, whereas the models developed by proponents of the social hypothesis typically target population variation over time. We also note the potential for applications outside the sapiens line. For example, one might apply the framework to help impose order on and assess various hypotheses regarding Neanderthal extinction. (For a recent theoretical overview on this topic, see Meneganzin and Currie (2022).)

But in addition to "zooming in" on debates about particular behavioural transitions, we might also "zoom out." In particular, the framework developed here might be used to assess certain single-factor explanations of human uniqueness. Put simply, these are accounts of the form "X made us human," for some X, and are surprisingly common. Recent prominent examples include cooperative breeding (Burkart & van Schaik, 2016; Hrdy, 2011); the domestication of fire (Wrangham, 2010); shared intentionality (Tomasello, 2010); pair-bonding (Chapais, 2009); and weapons (Bingham & Souza, 2009). Clearly, none of these theorists means to claim that understanding the role played by their preferred causal factor explains the *whole* of the evolution of human uniqueness. What exactly *is* meant, then?

Setting aside the merely pragmatic, we are inclined to think that such theorists may (tacitly) have in mind an actual difference making claim of some kind. The thought would be something like this: at some point in our evolutionary past, hominins still very much fell within a range of variation considered "normal" for great apes. Then something happened that put us on the human uniqueness trajectory. That trajectory itself has no doubt been highly causally complex, but perhaps its ultimate origins were not. Perhaps, that is, there was just a single actual difference making cause that kicked things off—a single cause variable that actually varied between us and other great apes at the start of this trajectory which explains why we wound up on this path and they didn't. That is indeed possible. But it is also possible that there never was just a single actual difference maker; that instead, humans and other great apes actually differed with respect to several causally relevant factors at the start of this process.

In closing: to us, it is primarily the expanded menu of causal concepts, along with how interventionists seek to ground these concepts in *objective* features of the world, that serves to distinguish interventionism from other frameworks that are sometimes used by human evolutionary researchers to think about causation, such as directed acyclic graphs (DAGs), structural equation modelling (SEM), and the Rubin causal model (RCM) (aka the potential outcomes framework). To the best of our understanding, the latter are more concerned with practical issues of how to infer causal relationships from (inherently noisy) data, as well as how to quantitatively measure the contributions made by various causes to a given effect. These are, of course, matters of great importance. But they are also largely complementary with the more overtly philosophical aims of interventionism. It is thus not surprising that interventionists draw on a variety of formal tools and ideas from these other frameworks in representing and reasoning about causal relations (see, e.g., Woodward 2003, §2.2-3 for an in-depth discussion of the role of directed causal graphs and structural equations in interventionist thinking). In terms of the specific issues we have discussed here: some of the measurement-theoretic tools developed by these other frameworks should be useful for measuring causal specificity in the interventionist sense. The notion of causal effect belonging to the RCM or potential outcomes framework can be plausibly used to estimate the degree of causal specificity of an actual difference maker, for example. For while it is impossible to compare outcomes in any individual case (an individual cannot simultaneously take and not take a pill, say), we can use an average of the distribution of outcomes over a population—some of whom have taken the pill and some of whom have not—to approximate the causal effect of taking a pill. There are significant methodological challenges when it comes to applying this approach to the historical sciences including human evolution, but researchers have made progress in overcoming these challenges (Lonati et al., 2024). If, as we have suggested, some debates in human evolution concern whether or not a cause is a specific actual difference maker, then the potential outcomes should be of use in helping us progress those debates.

The approach we have developed in this article strikes us as well-suited to clearly formulating various questions and challenges facing researchers in the field of human evolution, and in so doing, helping us to solve them. Much interesting work remains to be done in applying this framework to such issues.

References

Arbib, M. A. (2011). From Mirror Neurons to Complex Imitation in the Evolution of Language and Tool Use. *Annual Review of Anthropology*, *40*(1), 257–273. https://doi.org/10.1146/annurev-anthro-081309-145722

Bandini, E., Motes-Rodrigo, A., Steele, M. P., Rutz, C., & Tennie, C. (2020). Examining the mechanisms underlying the acquisition of animal tool behaviour. *Biology Letters*, *16*(6), 20200122. https://doi.org/10.1098/rsbl.2020.0122

Bandini, E., & Tennie, C. (2018). Naive, captive long-tailed macaques (Macaca fascicularis fascicularis) fail to individually and socially learn pound-hammering, a tool-use behaviour. *Royal Society Open Science*, *5*(5), 171826. https://doi.org/10.1098/rsos.171826

Bergström, A., Stringer, C., Hajdinjak, M., Scerri, E. M. L., & Skoglund, P. (2021). Origins of modern human ancestry. *Nature*, *590*(7845), 229–237. https://doi.org/10.1038/s41586-021-03244-5

Berwick, R. C., & Chomsky, N. (2016). *Why Only Us: Language and Evolution*. MIT Press.

Berwick, R. C., & Chomsky, N. (2019). All or nothing: No half-Merge and the evolution of syntax. *PLOS Biology*, *17*(11), e3000539. https://doi.org/10.1371/journal.pbio.3000539

Bingham, P. M., & Souza, J. (2009). *Death from a Distance and the Birth of a Humane Universe: Human Evolution, Behavior, History, and Your Future*. Booksurge Publishing.

Boyd, R. (2017). *A Different Kind of Animal: How Culture Transformed Our Species*. Princeton University Press.

Burkart, J. M., Hrdy, S. B., & Schaik, C. P. V. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *18*(5), 175–186. https://doi.org/10.1002/evan.20222

Burkart, J. M., & van Schaik, C. P. (2016). Revisiting the consequences of cooperative breeding. *Journal of Zoology*, *299*(2), 77–83. https://doi.org/10.1111/jzo.12322

Chapais, B. (2009). *Primeval kinship: How pair-bonding gave birth to human society*. Harvard University Press.

Cieri, R. L., Churchill, S. E., Franciscus, R. G., Tan, J., & Hare, B. (2014). Craniofacial Feminization, Social Tolerance, and the Origins of Behavioral Modernity. *Current Anthropology*, *55*(4), 419–443. https://doi.org/10.1086/677209

Cole, J. (2016). Accessing Hominin Cognition: Language and Social Signaling in the Lower to Middle Palaeolithic. In *Cognitive Models in Palaeolithic Archaeology* (pp. 157–195). Oxford University Press.

Cole, J. (2019). Knapping in the Dark: Stone Tools and a Theory of Mind. In *Squeezing Minds from Stones: Cognitive Archaeology and the Evolution of the Human Mind* (pp. 355–375). Oxford University Press. https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190854614.001.0001/oso-9780190854614-chapter-18

Collard, M., Buchanan, B., & O'Brien, M. J. (2013). Population Size as an Explanation for Patterns in the Paleolithic Archaeological Record: More Caution Is Needed. *Current Anthropology*, *54*(S8), S388–S396. https://doi.org/10.1086/673881

Collard, M., Kemery, M., & Banks, S. (2005). Causes of Toolkit Variation Among Hunter-Gatherers: A Test of Four Competing Hypotheses. *Canadian Journal of Archaeology / Journal Canadien d'Archéologie*, *29*(1), 1–19. JSTOR.

Collard, M., Vaesen, K., Cosgrove, R., & Roebroeks, W. (2016). The empirical case against the 'demographic turn' in Palaeolithic archaeology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1698), 20150242. https://doi.org/10.1098/rstb.2015.0242

Coolidge, F. L., & Wynn, T. (2018). *The Rise of Homo Sapiens: The Evolution of Modern Thinking* (2nd ed.). Oxford University Press.

de la Torre, I. (2016). The origins of the Acheulean: Past and present perspectives on a major transition in human evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1698), 20150245. https://doi.org/10.1098/rstb.2015.0245

Hausman, D. M., & Woodward, J. (1999). Independence, Invariance and the Causal Markov Condition. *The British Journal for the Philosophy of Science*, *50*(4), 521–583. https://doi.org/10.1093/bjps/50.4.521

Henrich, J. (2015). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press.

Hernan, M. A., & Robins, J. M. (2023). *Causal Inference: What If*. CRC Press.

Heyes, C. (2018). *Cognitive Gadgets: The Cultural Evolution of Thinking*. Harvard University Press.

Hiscock, P., & O'Connor, S. (2006). An Australian perspective on modern behaviour and artefact assemblages. *Before Farming*, *2006*(2), 1–10. https://doi.org/10.3828/bfarm.2006.2.4

Hrdy, S. B. (2011). *Mothers and Others: The Evolutionary Origins of Mutual Understanding*. Belknap Press.

Li, H., and R. Durbin (2011). "Inference of Human Population History from Individual Whole-Genome Sequences." Nature 475 (7357): 493.

Lonati, S., Lalive, R., & Efferson, C. (2024). Identifying culture as cause: Challenges and opportunities. *Evolutionary Human Sciences*, *6*, e9. https://doi.org/10.1017/ehs.2023.35

Mcbrearty, S., & Brooks, A. S. (2000). The revolution that wasn't: A new interpretation of the origin of modern human behavior. *Journal of Human Evolution*, *39*(5), 453–563. https://doi.org/10.1006/jhev.2000.0435

Mellars, P. (2005). The impossible coincidence. A single-species model for the origins of modern human behavior in Europe. *Evolutionary Anthropology: Issues, News, and Reviews*, *14*(1), 12–27. https://doi.org/10.1002/evan.20037

Mellars, P., & Stringer, C. (1989). *The Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans*. Princeton University Press.

Meneganzin, A., & Currie, A. (2022). Not by demography alone: Neanderthal extinction and null hypotheses in paleoanthropological explanation.

Mithen, S. J. (1996). *The prehistory of the mind: The cognitive origins of art,religion and science* (1st paperback). Thames and Hudson.

Muthukrishna, M., & Henrich, J. (2016). Innovation in the collective brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1690), 20150192. https://doi.org/10.1098/rstb.2015.0192

Nowell, A. (2010). Defining Behavioral Modernity in the Context of Neandertal and Anatomically Modern Human Populations. *Annual Review of Anthropology*, *39*(Volume 39, 2010), 437–452. https://doi.org/10.1146/annurev.anthro.012809.105113

Paddayya, M. D. P., Ceri Shipton, K. (2004). Life and Mind in the Acheulean: A Case Study from India. In *Hominid Individual in Context*. Routledge.

Powell, A., Shennan, S., & Thomas, M. G. (2009). Late Pleistocene Demography and the Appearance of Modern Human Behavior. *Science*, *324*(5932), 1298–1301. JSTOR.

Premo, L. S., & Kuhn, S. L. (2010). Modeling Effects of Local Extinctions on Culture Change and Diversity in the Paleolithic. *PLOS ONE*, *5*(12), e15582. https://doi.org/10.1371/journal.pone.0015582

Richerson, P. J., & Boyd, R. (2008). *Not By Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press.

Richerson, P. J., & Boyd, R. (2013). Rethinking paleoanthropology: A World queerer than we supposed. In G. Hatfield & H. Pittman (Eds.), *Evolution of Mind, Brain, and Culture* (1–Book, Section, pp. 263–302). University of Pennsylvania Press.

Schiffels, S., and R. Durbin (2014). "Inferring Human Population Size and Separation History from Multiple Genome Sequences." Nature Genetics 46 (8): 919–925.

Schlebusch, C. M., Skoglund, P., Sjödin, P., Gattepaille, L. M., Hernandez, D., Jay, F., Li, S., De Jongh, M., Singleton, A., Blum, M. G. B., Soodyall, H., & Jakobsson, M. (2012). Genomic Variation in Seven Khoe-San Groups Reveals Adaptation and Complex African History. *Science*, *338*(6105), 374–379. https://doi.org/10.1126/science.1227721

Shultziner, D., Stevens, T., Stevens, M., Stewart, B. A., Hannagan, R. J., & Saltini-Semerari, G. (2010). The causes and scope of political egalitarianism during the Last Glacial: A multi-disciplinary perspective. *Biology & Philosophy*, *25*(3), 319–346. https://doi.org/10.1007/s10539-010-9196-4

Sterelny, K. (2014). A Paleolithic Reciprocation Crisis: Symbols, Signals, and Norms. *Biological Theory*, *9*(1), 65–77. https://doi.org/10.1007/s13752-013-0143-x.

Sterelny, K. (2021b). The origins of multi-level society. *Topoi*, *40*(1), 207-220.

Sterelny, K. (2021). *The Pleistocene social contract: Culture and cooperation in human evolution*. Oxford University Press.

Stout, D., & Chaminade, T. (2012). Stone tools, language and the brain in human evolution. *Philosophical Transactions: Biological Sciences*, *367*(1585), 75–87. JSTOR.
Stout, D., Chaminade, T., Apel, J., Shafti, A., & Faisal, A. A. (2021). The measurement, evolution, and neural representation of action grammars of human behavior. *Scientific Reports*, *11*(1), 13720. https://doi.org/10.1038/s41598-021-92992-5
Straus, L. G. (2000). A Quarter-Century of Research on the Solutrean of Vasco-Cantabria, Iberia and beyond. *Journal of Anthropological Research*, *56*(1), 39–58.
Tattersall, I. (2016). A tentative framework for the acquisition of language and modern human cognition. *Journal of Anthropological Sciences = Rivista Di Antropologia: JASS*, *94*, 157–166. https://doi.org/10.4436/JASS.94030
Tennie, C. (2023). The Earliest Tools and Cultures of Hominins. In J. J. Tehrani, J. Kendal, & R. Kendal (Eds.), *The Oxford Handbook of Cultural Evolution* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198869252.013.33
Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and brain sciences*, *28*(5), 675-691.
Tomasello, M. (2010). *Origins of Human Communication*. MIT Press.
Torrence, R. (1983). Time budgeting and hunter-gatherer technology. In G. Baily (Ed.), *Hunter-gatherer economy in prehistory: A European perspective* (pp. 11–22). Cambridge University Press.
Torrence, R. (1989). Retooling: Towards a behavioral theory of stone tools. In R. Torrence (Ed.), *Time, energy, and stone tools: Vol. New directions in archaeology* (pp. 57–66). Cambridge University Press.
Torrence, R. (2001). Hunter-gatherer technology: Macro- and microscale approaches. In C. Panter-Brick, R. H. Layton, & P. Rowley-Conwy (Eds.), *Hunter-gatherers: An interdisciplinary perspective* (pp. 73–98). Cambridge University Press.
Waddington, C. H. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, *150*(3811), 563-565.
Waters, C. K. (2007). Causes That Make a Difference. *The Journal of Philosophy*, *104*(11), 551–579. https://doi.org/10.5840/jphil2007104111
Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.
Woodward, J. (2010). Causation in biology: Stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, *25*(3), 287–318. https://doi.org/10.1007/s10539-010-9200-z
Woodward, J. (Forthcoming). Reflections on Some Strategies for Causal Inference in Psychiatry. In K. Kendler, P. Zachar, & J. Parnas (Eds.), *Philosophical Issues in Psychiatry VI*. Cambridge University Press.
Wrangham, R. (2010). *Catching Fire: How Cooking Made Us Human*. Profile Books.
Wynn, T., & Coolidge, F. L. (2004). The expert Neandertal mind. *Journal of Human Evolution*, *46*(4), 467–487. https://doi.org/10.1016/j.jhevol.2004.01.005