

The Evolution of Cultural Complexity

Not by the Treadmill Alone

by Claes Andersson and Dwight Read

Among the drivers and constraints on the evolution of complex hominin culture that have been proposed throughout the years, demographic factors have been particularly persistent, and they have recently again come to gain traction in the literature in the shape of the so-called treadmill model. The treadmill model connects cultural complexity to group size via a need to constantly “outrun a treadmill of cultural loss,” whose backward motion is caused by errors in culture transmission. The entrenchment of the treadmill explanation of cultural complexity, however, takes place against a background of critiques of the model and the presence of other explanatory propositions. This creates a need for deentrenchment: wider integration, elaboration, and critique of the premises of the treadmill model and the evidence advanced to validate it. We begin by reviewing the treadmill model, making an assessment of its current status, and then moving on to a more synthetic proposition by placing the model into the context of other models addressing the elaboration of cultural complexity. We end by considering the broader implications for the study of the evolution of culture and of human behavior to be gained from more integrated modeling of the various factors affecting cultural complexity.

Introduction

The question of what drove and shaped hominin dependence on increasingly complex cultural systems is clearly a key piece for understanding behavioral changes that have emerged during human evolution, and it has long been a subject of considerable speculation and debate (e.g., Cohen 1985; Johnson and Earle 2000). One potential driver that has received much attention, both historically and over the past decade, is a demographic factor leading to an evolutionary pattern going from small and weakly interacting groups, confined to a narrow range of habitats, to increasingly interconnected groups capable of thriving in even the most extreme of habitats across the planet (e.g., Gamble 2007, 2010). Understanding this pattern and how it relates to increased cultural complexity provides the focus for our synthetic review.

Two influential types of causal connection between demographic change and increased cultural complexity have been posited. We will refer to one as “the traditional connection” and to the other, more recently proposed, as a “new connection.”

The traditional connection is based on the interactions among groups and their environments, in particular with re-

gard to the means for obtaining food resources. The envisioned connection in this older scenario is between population pressure (e.g., Flannery 1976; Johnson and Earle 2000; Keeley 1988) and complexity. To sustain a larger population on a given piece of land, the available amount of resources must increase, and one important way to accomplish this is through more complex technology for resource extraction, storage, and processing. Cultural organization and dynamics are not explicitly modeled but are instead subsumed under optimality and rationality assumptions. In Flannery’s (1969) broad-spectrum revolution model, for example, the interplay between demography and environment is seen as giving rise to a geographically framed dynamic of domination and encroachment, whereby groups capable of extracting a broader spectrum of resources can expand in size and thereby gain a competitive edge over groups exploiting a narrower spectrum of resources (Read 1987). This potentially creates a geographical bootstrapping dynamic in which competition perpetually renews the pressure for increasing resource output, thereby pushing for increased complexity in social organization and tool technology as a way to intensify, broaden, and increase the effectiveness of the extraction and use of resources (Read and LeBlanc 2003; Stiner and Kuhn 2006; Stiner and Munro 2000).

The “new connection,” though, has no obvious continuity with the traditional connection outlined above. It belongs, instead, to a broader effort to identify endogenous drivers and constraints on culture that emerge from social interaction and institutions. Linking demography and cultural complexity in this manner was first proposed by Shennan (2001), but the model that has come to be the most influential was put for-

Claes Andersson is an Associate in Physical Resource Theory in the Department of Energy and Environment at Chalmers University of Technology (412 96 Göteborg, Sweden [claeand@chalmers.se]). **Dwight Read** is Distinguished Professor Emeritus in the Department of Anthropology of the University of California, Los Angeles (Los Angeles, California 90095, USA [dread@anthro.ucla.edu]). This paper was submitted 2 XII 14, accepted 18 V 15, and electronically published 29 IV 16.

ward in a seminal paper by Henrich (2004). In brief, Henrich argues that imitation is imperfect and necessarily involves an immediate loss of information that must be dynamically compensated for by the creativity of individuals to maintain the long-term persistence of cultural systems. This process has been described as a “treadmill of cultural loss” (Kline and Boyd 2010) against which a population must constantly run even to stay in the same place. The larger the population, Henrich argues, the more likely that highly creative individuals will be part of the population, thereby increasing the likelihood of augmenting and elaborating transmitted skills even beyond what is needed to compensate for information loss. According to Henrich (2004), a larger population can “run faster” to counteract the treadmill effect and maintain skillfulness in more complex cultural traits.

Since 2004, additional theoretical, empirical, and experimental work, aimed at developing and supporting the treadmill model, has been published. The original model, along with its subsequent derivations, has been widely cited as valid by an increasingly public audience.¹ For example, recently in *Nature*, Richerson (2013:351) confidently states—as the title of a commentary on a recent experiment aimed at testing the treadmill model by Derex et al. (2013)—that “group size determines cultural complexity” (see also Bell 2014, 2015; Chaisson 2014; Richerson et al. 2016). The hypothesis that demography is the main causal factor behind the evolution of cultural complexity in humans is now prominently treated as an established fact, and the treadmill model has become the canonical model.

But at the same time, the treadmill model has been vigorously contested, both theoretically and empirically, and with regard to both its premises and its predictions, since its inception (e.g., Andersson and Read 2014; Collard, Kemery, and Banks 2005; Collard et al. 2013a, 2013b; O’Brien and Bentley 2011; Querbes, Vaesen, and Houkes 2014; Read 2006, 2008, 2009, 2012b; Vaesen 2012a). There are also other proposed drivers and constraints on cultural complexity that are being proposed, such as fidelity (e.g., Andersson 2011, 2013; Claidiere and Sperber 2010; Enquist et al. 2010; Tennie, Call, and Tomasello 2009; Tomasello 1991), pedagogy (e.g., Csibra and Gergely 2009, 2011; Sterelny 2011; Tostevin 2012), and differential complexity as an adaptive response to environmental variables (Collard et al. 2013a; Read 2008; Torrence 1983, 1989).

In sum, the widespread interest in the treadmill model, its increasing entrenchment despite the unresolved criticisms that it faces, and the wide range of other hypotheses addressing the same question that are being vetted create a need for a critical review of the treadmill model and a synthesis of the disparate viewpoints, all placed into a broader context. This is what we aim to provide in this article.

We begin by describing the basic treadmill model as introduced by Henrich (2004). Our aim is to do this in a widely accessible way and to clarify some aspects that have been sources of consequential confusion. Next, in our review of the debates surrounding the treadmill model, we find it useful to partition the debate into two aspects concerning (i) the predictions and (ii) the premises of the treadmill model. In our review of the former, we identify three subdebates about (a) historical case studies, (b) data on metapopulations of forager groups, and (c) experimental tests.

We then differentiate between two interpretations of the role of the treadmill model in explanations of variability in cultural complexity. The strong interpretation holds that the treadmill model, in some form, represents the prime mover of cultural complexity. The weak interpretation holds that it represents one among several causal factors.

Overall, the weight of the evidence tells us that the strong interpretation of the treadmill model should today be viewed as untenable, though the model may be relevant under a weaker interpretation. We outline such a role in a wider synthesis that includes the traditional connection and other drivers and constraints on cultural complexity previously proposed in the literature. Not least, we call for endogenous and exogenous factors to be used in combination, pointing to similar directions in modern evolutionary theory. The latter points strongly to limitations in seeking evolutionary explanations based on single causal factors and on a single level of organization.

Treadmill Model

The treadmill model springs from the context of the dual-inheritance theory of evolution (DIT; Boyd and Richerson 1985; Cavalli-Sforza and Feldman 1981; Mesoudi 2011b), which is a central element in the recent attention being paid to endogenous cultural explanations. Although this new demography-complexity link was first introduced by Shennan (2001), it is not his model, but a different model introduced briefly by Henrich and Boyd (2002) and then more completely and anthropologically contextualized by Henrich (2004), that has come to take center stage in theoretical and empirical research on this new and different version of the demography-complexity link. Henrich’s (2004) model can be divided into two major parts: (i) a formal analytical model and (ii) a string of informal extrapolations that take us from the analytical model to the predictions of the treadmill model as a whole (fig. 1).

The formal model depicts a population of size N with socially interacting individuals that maintain a “cultural trait” that is adaptive to the group in some manner. Following Henrich (2004), we refer to the cultural trait (e.g., technology or practice) as “the skill” and to how well the skill is performed as “skillfulness,” denoted by z in the treadmill model.² What is

1. With over 350 citations of Henrich’s original article (Henrich 2004) and just over 300 citations to the article by Powell, Shennan, and Thomas (2009) identified by Google Scholar.

2. It should be noted that Henrich (2004) is not consistent in making this distinction; for example, in the abstract and on pages 198, 203–204, and 207–209, “skill” represents a cultural trait, but on pages 200–202,

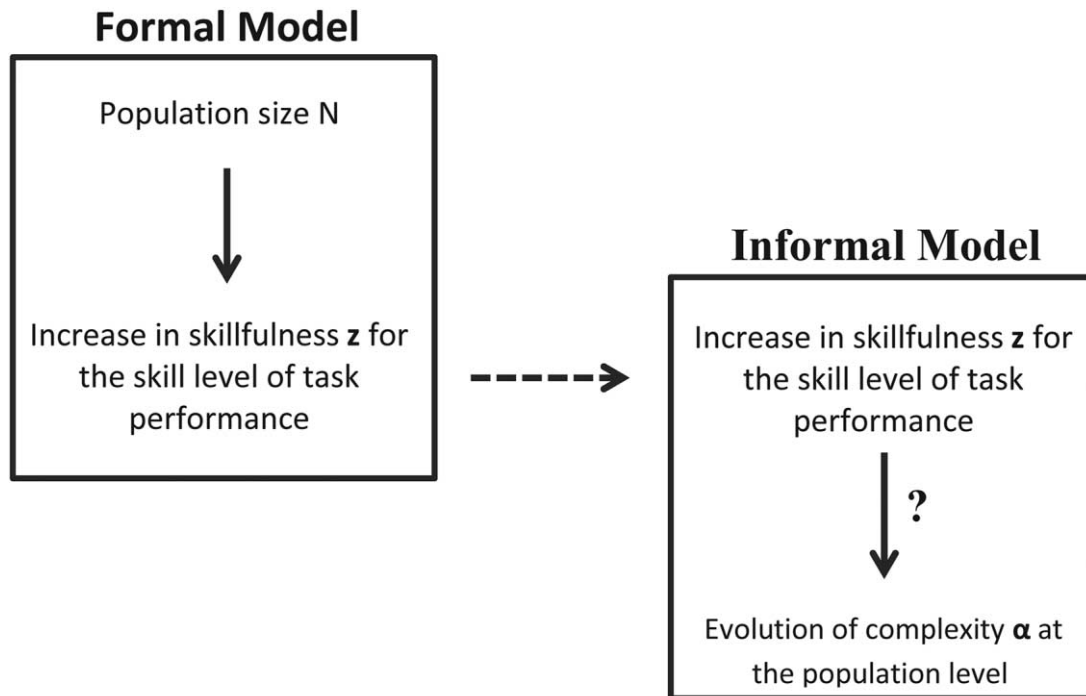


Figure 1. The chain of explanation between population size and complexity, via the concept of a level of skillfulness with which a skill is exercised. It is worth noting that the formal model only concerns the expected change in average skillfulness across one round of imitation. The rest of the model, most notably the link between changes in skillfulness and the complexity of skills in a population (which provides the full link between population and complexity) consists of an informal, incomplete, evolutionary extension of the formal model, as indicated by the question mark.

depicted formally is the evolution of skillfulness z with which the skill is performed. If skillfulness drops too low, the model assumes that the skill will be useless and abandoned.

Henrich begins by assuming the skill is transmitted obliquely³ in the population through imitation, which is invoked in the broad sense that naive individuals reconstitute the transmitted skill from interaction with selected role models. Selection here only implicitly refers to any economic or other consequences of using the skill in question. Explicitly, it refers to the process by which naive individuals select highly skillful role models to imitate, introducing what is referred to in the DIT literature as a “skill bias.” The imitation dynamic involves, then, two steps: (i) locating a role model and (ii) imitating the selected role model.

Henrich simplifies the first step by assuming that naive individuals always find the most skillful role model in the population to imitate. This motivates his use of the Gumbel dis-

tribution (which is an extreme value distribution, the choice of which, moreover, simplifies the mathematical modeling) to model the statistical frequency distribution for the most skillful person in a population of size N .⁴ This also provides a link between population size and skillfulness (i.e., completing the formal model in the chain of explanation; fig. 1). The larger the population, the greater the skillfulness we expect to find for the most skillful person in the population.

The second step incorporates two types of variability in the imitation process: (i) that the quality of the imitations is, on average, worse than the original target, because inevitable differences between copy and original are more likely to be deleterious than beneficial; (ii) that imitation outcomes vary in quality, because imitators, like role models, are differentially gifted in terms of creativity and learning.

We may schematically illustrate the acquisition of a skill by a naive imitator (obtaining skillfulness z') from a role model (with skillfulness z) by the equation

$$z' = z - \alpha + e. \quad (1)$$

when the formal model is presented, “skill” is measured by z and hence now denotes skillfulness.

3. That is, transmission between generations, but without a bias for transmission along family lines. Besides oblique transmission, the other relevant transmission modalities in this context are vertical transmission between parents and offspring and horizontal transmission between peers in the same age cohort.

4. Under this assumption, the frequency distribution for the skill level of the role models across a collection of populations of size N , with one role model selected from each population (assuming each population has the same parameter values), will converge to a Gumbel distribution as N increases.

Here, α represents the first type of variability and quantifies how hard the skill is to imitate. What makes a skill hard to learn is identified as “complexity,” and so α is also interpreted as a measure of skill complexity, and it is invoked in units of z . The link between the mechanics, behavior, and interpretations of the treadmill model pivots on this dual nature of α -as-error and α -as-complexity. Like Henrich (2004), we will refer to α both as “the average error in imitation” and as “complexity” interchangeably. The term e corresponds to the second type of variability, with values from the Gumbel distribution.

The imitation process is compactly modeled using a Gumbel distribution with location $z - \alpha$ and scale β , which measures the degree of variability in the e values (fig. 2). Crucially, for the functioning of the model, under some parameter regimes, it becomes likely that at least one imitation will be as good as, or better than, the original; in figure 2, this likelihood is illustrated by the green portion of the probability density function where $z' > z$. When this is the case, the population can outrun the treadmill. In the informal model, this is interpreted as meaning that cumulative cultural evolution is possible.

Henrich now uses the Price equation (Price 1970) to model the change in average skillfulness, $\Delta\bar{z}$, in one step of an evolutionary process governed in the manner discussed above and obtains

$$\Delta\bar{z} = -\alpha + \beta(\varepsilon + \ln N), \quad (2)$$

(see eq. [2] in Henrich 2004). Keeping the skill-specific parameters α and β fixed in equation (2) allows investigation of how $\Delta\bar{z}$ changes as a function of population size N . Equation (2) implies that, for N smaller than a threshold value $N_{\min}(\alpha)$, $\Delta\bar{z}$ turns negative.⁵ $N_{\min}(\alpha)$ then marks the smallest population for which a skill of complexity α can be sustained. Below $N_{\min}(\alpha)$, the skill will undergo what Henrich refers to as “maladaptive loss.” In other words, positive selection for skillfulness does not help; the level of skillfulness will drop, and the skill will be lost.

But the main prediction of Henrich (2004) is not about decreases in skillfulness z ; it is about decreases in complexity α . It is important to note, because it has been a source of confusion, (i) that Henrich’s formal model does not concern changes in α , (ii) that z cannot be interpreted as complexity without seriously compromising the logic of the treadmill model, (iii) that Henrich (2004) hints at, but does not provide, any model for changes in complexity α in individual skills, and (iv) that what Henrich (2004) does propose is an informal model for decreases in α with decreases in population size (see “informal model” depicted in our fig. 1).

5. Henrich (2004) refers to complexity as α/β . For brevity, and following, for example, Powell, Shennan, and Thomas (2009) and Vaesen (2012a), we refer to complexity simply as α . In figure 3, we fix $\beta = 1$ and vary α .

The link between losses in skillfulness z in individual skills and losses in complexity α in the skill population derives from the preferential elimination of skills corresponding to high complexity α in a population of skills in the group (fig. 3). This also means that, in the shift from the formal model of z to the informal model of α (fig. 1), there is also a shift from a single skill to a population of skills. For this model of the loss of complexity—which is entirely informal in Henrich (2004), yet delivers the main prediction of the treadmill model in its entirety—to work, we need to lay bare two implicit assumptions and arguments.

The first assumption is necessary and has to do with the mix of skills that groups possess. We may derive from the formal model that, for a population of size N , there is a threshold level, $\alpha_{\max}(N)$, above which skills would need a larger population size to survive; see figure 2 in Henrich (2004). To tie this to maladaptive losses of skills (not skillfulness), we need to assume that cultures will contain a mix of skills with different α values lower than $\alpha_{\max}(N)$. These should be distributed all the way up to the $\alpha_{\max}(N)$ threshold value—or at least close to it. The culture in question will then, with a population decrease from N^{high} to N^{low} , preferentially lose their most complex skills by the process depicted by the formal model. They will be left with a mix containing skills with α values below a new and lower threshold $\alpha_{\max}(N^{\text{low}})$ (see fig. 3). Unless this assumption holds reasonably well, a population decrease would be likely to have no effect. We shall see later that this is a strong assumption, depending on how we imagine that the population is structured (i.e., what the population size N corresponds to).

The second assumption is made to maintain analytical tractability in the formal model, and it is also a source of considerable confusion in the literature. It is the assumption that z corresponds strictly to level of skillfulness and that α corresponds strictly to the complexity of the skill or, interchangeably, to how hard it is to learn.⁶

The problem with this is that, contrary to this assumption, z and α are not made distinct in verbal descriptions of the processes that the model represents. This is probably due to the simple fact that it is hard to imagine (let alone, as we will argue, design an experimental rendition of) a process where skillfulness and complexity are, in general, independent. We would have to imagine an innovation process wherein the level of skillfulness changes but where the complexity of the skill strictly does not change. For example, under this assumption, a less skillful version of a bone point (lower z) could not be a simpler bone point (lower α). It must be exactly the same bone point—precisely as hard or easy to imitate—only less skillfully made.

6. Henrich (2004) stresses that α and β are not intrinsic to the skill itself but to both the skill and human cognition. But since human cognition is, in practice, taken to be constant, this means that it varies only with the nature of the skill.

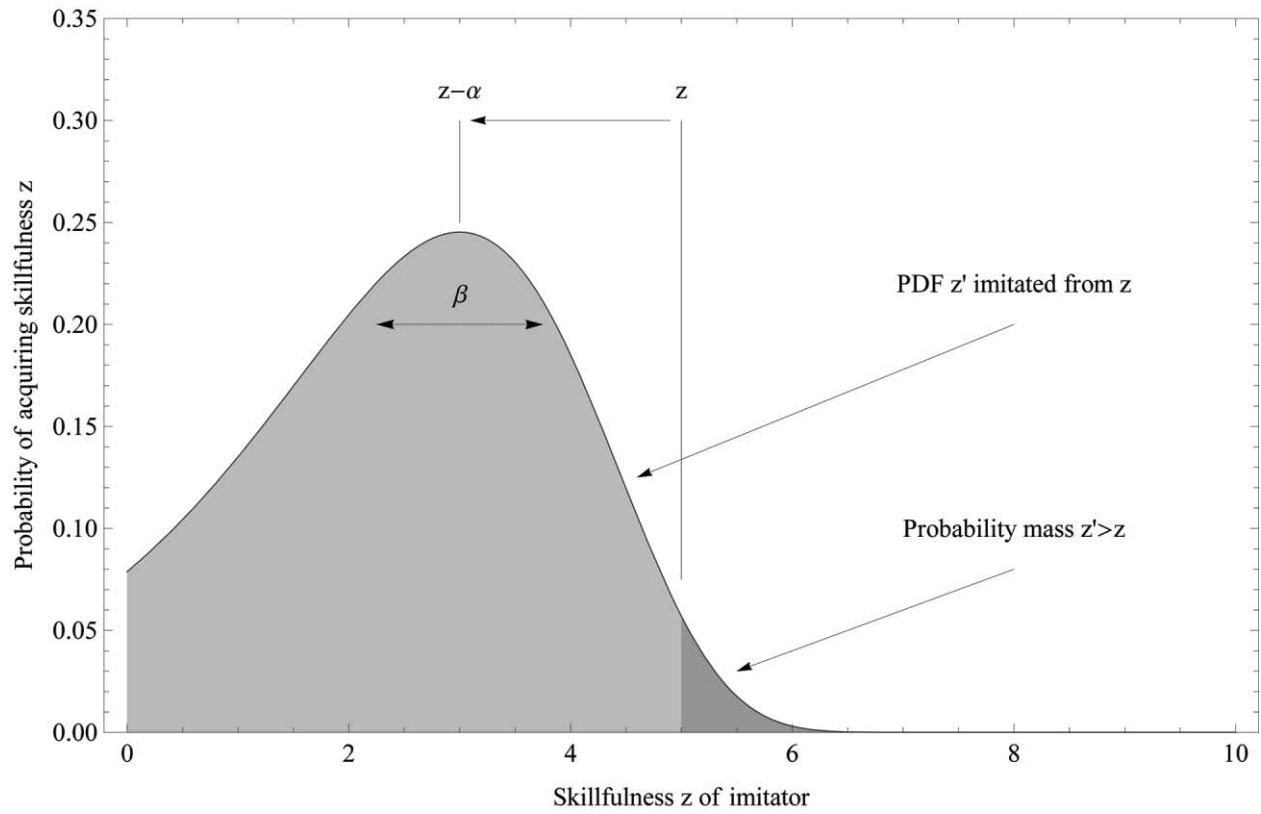


Figure 2. Mechanism of skill acquisition in Henrich (2004). We illustrate the draw of new skillfulness value z' for a naive learner imitating a role model with skillfulness z . The probability density function (PDF) of a Gumbel distribution, with location $z - \alpha$ and scale β , with the size of the probability mass for the learner besting the role model (i.e., drawing $z' > z$) indicated in dark gray. A color version of this figure is available online.

This degeneracy between z and α is not a minor point, unimportant in relation to the overall process. First of all, if α in individual skills were to change as part of the process by which z is imagined to change, then the skill-level dynamics would change substantially. Complexity α would be dynamical, and skill deterioration could, for example, lead to a skill that would be easier to learn (i.e., lower α), which could “save it” by bringing it back into the cumulative regime again, leading to an equilibrium between α and z , as long as minimal requirements on how simple the skill can be and remain effective are met. But perhaps even more serious (the previous scenarios would still predict a reduction in complexity) is the fact that it is hard both to extend and to get straight answers from a model whose two most central concepts are degenerate in unknown and unexamined ways. We will repeatedly have reasons to return to how this confusion impacts research on the treadmill model negatively in several ways (in interpretations, in elaborations, and in experimental testing) throughout our review.

Let us briefly illustrate the downstream confusion that this has caused by considering the model proposed by Mesoudi (2011a). Mesoudi argues that Henrich’s (2004) model does not consider the possibility that more complex skills can be

more costly to transmit and maintain. The result, he argues, is that Henrich’s model predicts an indefinite increase in skill complexity in the cumulative regime where $\Delta \bar{z} > 0$. His solution is to attach a cost to z such that $\Delta \bar{z}$ will tend to 0 over many generations, leading to a stable equilibrium in z . But, as we have explained, z does not correspond to complexity in Henrich’s model. Mesoudi here falls victim to the poor (but arguably convenient) delineation between z and α , and it is clear from his argument that he is describing increases in complexity to motivate why increases in z (which he erroneously interprets as complexity) would lead to higher complexity costs.

Testing the Predictions of the Treadmill Model

Predictions about empirical observations derived from the treadmill model have been used in attempts both to validate the model and to disprove it. Several data sets have figured in this debate: (i) archaeological evidence regarding changes in the tool assemblages of the Tasmanian hunter-gatherers around the time (8,000 BP) when they were isolated from mainland Australia, (ii) data on variability in the complexity of hunter-gatherer tool assemblages (Oswalt 1976), (iii) data

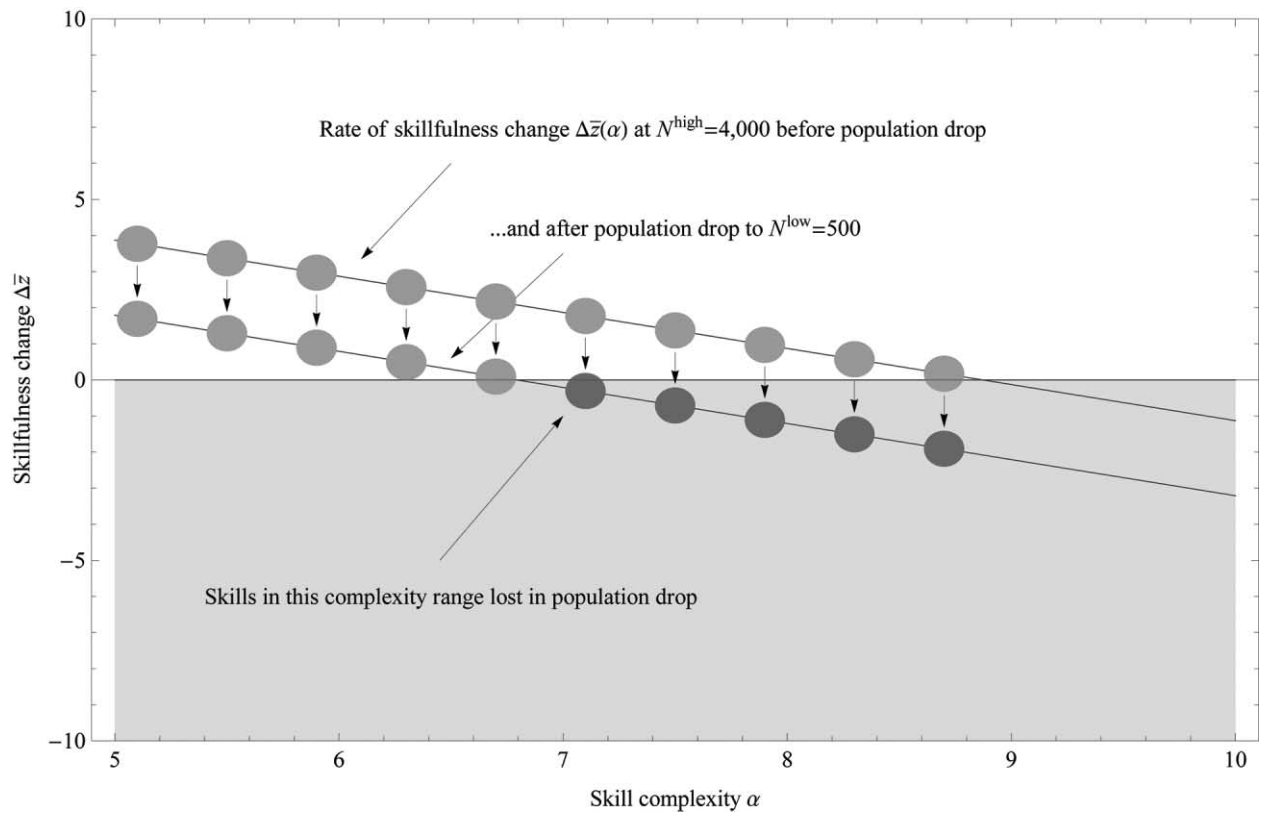


Figure 3. Mechanism of loss of cultural complexity α at the population level via loss of skillfulness z on the level of skills. Decreasing the population from $N = 4,000$ to $N = 500$, an interval of skills with high values of α going from positive to negative values of $\Delta \bar{z}$ (see eq. [2] here and in Henrich 2004). This moves the skills from the cumulative to the maladaptive (red) regime, leading to their deterioration and abandonment by the process illustrated in Figure 2. It should be noted that the magnitude of $\Delta \bar{z}$ is not centrally important; apart from some statements about the speed of evolution, the overshadowing importance of $\Delta \bar{z}$ for the main thrust of the treadmill model is whether it is positive or negative. A color version of this figure is available online.

on tool complexity for Oceania Island subsistence fishing groups (Kline and Boyd 2010), and (iv) historical and modern data on Inuit tool complexity and population sizes (references in Read 2012b). Experiments have also been conducted to determine whether predictions regarding the number of imitators and tool complexity are observed under experimental conditions. Here we group the debate about predictions into three categories: (a) historical case studies, (b) data on metapopulations of forager groups, and (c) experimental tests.

Historical Case Studies

On the face of it, the Tasmanian case study appeared to be ideal for testing the model, because (i) the size of the potential group of imitators likely dropped substantially when Tasmania became isolated from mainland Australia, whereas (ii) their tool assemblage appears to have decreased in complexity. This is indicated by the disappearance (with the exception of one archaeological site) of the bone points used to make clothing before the rise in sea level around 8,000 BP.

The model's prediction of maladaptive loss also appeared to be validated by the fact that, at time of contact with Europeans, the Tasmanians were not wearing clothing despite experiencing a cold climate. These data had already been interpreted as indicating that the decrease in the artifact assemblage of the Tasmanians was due to change in population size (Jones 1971), and the treadmill model stepped in to provide the demographic rationale for this interpretation.

If the model is valid, the primary predictions that should be clearly validated are that (i) the lost skills should have been comparably complex (large α) and (ii) the loss should be maladaptive rather than caused by external selective forces. In addition, one should expect to see a deterioration in skillfulness (z), although that is a weaker prediction, because the time scale for deterioration is hard to determine.

Counterarguments have focused on six points:

1. Despite claims to the contrary, the only documented loss of a tool is that of bone points (Hiscock 2008).
2. The bone points were not particularly sophisticated. They have been described as "a low-level innovation, are easy to

make” (Mulvaney and Kamminga 1999:355) using a simple technology (Buc 2010; Newcomer 1974; Tyzzer 1936) that is easy to imitate. The Tasmanians, moreover, made simple, not complex, clothing (Gilligan 2007a).

3. Lithic artifacts continually made by the Tasmanians were more complex than the bone points that were lost.
4. There is no evidence that the quality of the bone points declined after Tasmania was isolated, and the points after 8,000 BP may have been used to make nets (Bowdler and Lourandos 1982) or box traps (Hiscock 2008).
5. There are environmental factors that explain why Tasmanians could have stopped making bone points and clothing. With but one exception, where the points may have been used for making nets, bone points only occur during extremely cold periods (see fig. 2 in Read 2012b), when “simple clothing” (Gilligan 2007a) was made in response to environmental conditions. After the climate substantially ameliorated at the end of the last Ice Age, the need for clothing diminished (see Jones 1990). They continued, moreover, to make and wear simple skin cloaks during the colder parts of the year (Gilligan 2007a).

The fact that the Tasmanians abandoned fishing, despite an abundance of fish in the surrounding waters, has also been put forward as evidence of maladaptive losses. This assessment is, however, not without problems. A study of the historical Tasmania diet concluded that it was “considerably in excess of protein and greatly deficient in carbohydrates . . . the whole existence of this race was a permanent struggle to satisfy the craving of the body for carbohydrates . . . they were never able to provide a sufficient supply” (Noetling 1911:303). To make up for the shortage of plant carbohydrates, animal fat and bone marrow can be consumed to counteract the negative consequences of a diet that otherwise uses protein as a major caloric source (Speth and Spielman 1983:13). This may be the reason that sites prior to the Holocene in southwest Tasmania have thousands of wallaby and wombat bones broken open to obtain the marrow (Cosgrove and Allen 2001). The single fish species (*Pseudolabrus* species) found in some sites before around 3,800 BP was, however, a lean fish and would not have been a source of fat (Allen 1979 and references therein). A source of carbohydrates was the shellfish, exploited through historical times. Shellfish, such as the abalone and mollusks, exploited extensively by the Tasmanians, have 5%–10% carbohydrate by weight (Texler 1983) and so were an important source of carbohydrates in an otherwise carbohydrate-short diet. The Tasmanians did not consider fish to be edible, likely due to the fact that fish were neither needed for protein nor a source of carbohydrates; hence, the investment required to obtain fish may actually have been maladaptive.

But there is a wider conceptual and empirical problem hiding here. The strong focus on technology in cultural evolution research clearly stems from the fact that technological skills can be tracked archaeologically. But what if groups with low technological complexity compensate by maintaining com-

plex nontechnology skills? Taylor (2010) makes the case that cultures can be expected to pursue either of two trajectories with respect to investment in material technology. The first is the one that we tend to expect, where the functioning of the body is augmented by complex material technology. Reliance on material technology, however, also has the effect of entangling individuals in various requirements, such as obtaining and transporting raw material, maintaining and repairing artifacts, and dealing with the risk of technology failing (Hodder 2012). This indicates that, under certain circumstances, it might be more beneficial to go in a direction that minimizes the dependence on material technology, replacing it with nonmaterial skills instead. Taylor (2010) argues that both the Tasmanians and the Fuegians, with their simple tools, show clear evidence of having pursued such a trajectory in their development of cultural strategies for dealing with their environments. This would mean that their low technological complexity says little about the complexity of the skills they maintained in general.

Metapopulations

Oswalt’s (1976) data set has been used to show that there is a correlation between neither the census population size nor the population density (as a proxy measure for the number of imitators; see Read 2012b) and complexity of tool kits for hunter-gatherer societies located in regions covering all of the major environmental and ecological zones (Collard, Buchanan, and O’Brien 2013; Collard, Kemery, and Banks 2005; Collard et al. 2013a; Read 2006). Read (2008) has shown that over 95% of the variability in the complexity of hunter-gather implements relates to risk and mobility, the two variables that have long been used to account for the variability in the complexity of hunter-gatherer implements.

The main counterarguments to these data are twofold: (i) that the treadmill model refers to number of imitators, not the number of persons living together, and thus the effective sample size for computing statistical significance is less than the actual sample size; this is counter balanced by the fact that the sample size for computing statistical significance does not include those who are not involved in the imitation, such as children and the other sex for tasks done by a single sex; and (ii) that Oswalt’s data set focused primarily on Pacific Northwest hunting and gathering groups.

In reply, Collard et al. (2013a) and Collard, Buchanan, and O’Brien (2013) have expanded the data set to include additional hunter-gathering groups from outside the Pacific Northwest and have included measures of population size more obviously related to the interaction population size. The augmented data sets lead to the same conclusion: No relationship can be verified between population size, or interaction population size, and complexity of hunter-gatherer tool kits.

Oceanic island data set. Because of the absence of data on the interaction population size for hunter-gatherer groups,

Kline and Boyd (2010) developed a data set restricted to subsistence Oceanic Island fishers for whom at least a nominal measure of interaction with other groups was available. They showed statistically that risk measured by global measures, such as risk of typhoons, did not account for complexity measured by the number of different kinds of fishhooks. They also argued that, while the complexity of the fishhooks did not correlate significantly (5% level) with the presence other group interaction, it was almost significant and in the right direction.

Counterarguments (Read 2012*b*) have pointed out, for example, that Hawaii, one of their data points, did not practice subsistence fishing—which Kline and Boyd stipulated as a requirement for inclusion of groups in their data set—but made extensive use of fish farming. In addition, the data point for Hawaii is an obvious statistical outlier in a scatter gram plot of population size versus number of tool types, even when taking degree of interaction with other groups into account (see fig. 4 in Read 2012*b*). When Hawaii is removed from the data set, there is no longer even weak evidence relating fishing hook complexity to interaction with other groups. In addition, Kline and Boyd measured risk on a yearly basis, whereas the relationship between risk and complexity for hunter-gatherer groups is that between complexity and risk on each episode of hunting. Read (2012*b*) showed that when a measure of fishing risk for an episode of fishing, namely the risk of ocean currents during a fishing episode, was included, then the measure of contact with other groups drops out of the model. Thus, as is the case for the hunter-gatherer data, the risk model is a better predictor of the complexity of the fishing hooks than is the treadmill model.

Inuit data set. The Inuit made some of the most complex implements of any hunter-gather group, including a 36-part harpoon made by the Angmaksalik of eastern Greenland, a group of about 400 Inuit (see references in Read 2012*b*). The problem that the Inuit data pose for the treadmill model is that, even if we unrealistically assume that the total population of about 6,000 Inuit in all of Greenland⁷ was the “population of interacting social learners,” they still lacked the imitator population size that Henrich posits was needed for the Tasmanians to make even simple one-part bone point implements. A more realistic size of the population of interacting social learners for the Angmaksalik would be smaller than their census population size of about 400 persons (including women and children) due to their isolation from the Inuit in southern Greenland, an isolation that has been sufficient to maintain their genetic distinctiveness from them (see references in Read 2012*b*). In response to these and other data for the Inuit, Henrich now agrees that the complexity of the Inuit implements is due to their adapting to their environment and not due to the treadmill model (see “Supporting Material” in Read 2012*b*).

7. With local groups separated by thousands of kilometers of rugged coast line as one traverses from eastern to southern to western Greenland.

Experiments

Bringing hypotheses about cultural evolution to empirical test is challenging. Archaeological data are, in many ways, limited and biased, whereas ethnographic investigations are limited by the simple circumstance of the time scales that are involved. A recent and welcome addition to the repertoire, therefore, are experimental tests in which at least some of these limitations can be circumvented. We now briefly review some recent experimental work that has been brought to bear on the treadmill hypothesis.

Caldwell and Millen (2010) used the distance flown by a paper model airplane as a measure of performance and correlated it with group size after several rounds of imitation. The group members had the paper planes made in a previous round as models for the current round. The group size varied from 1 to 3 individuals. This study found no relationship between group size and performance of the paper airplanes.

Muthukrishna et al. (2014) have critiqued this experiment, suggesting that the model (i.e., the paper airplane) was too simple and could easily be imitated by anyone in a group. Hence, they argue, varying the group size from 1 to 3 would have little, if any, effect on group performance. Kempe and Mesoudi (2014) found difficulties with a different aspect of the experiment, namely that the individually made airplane designs made could not be easily integrated together to make a better design for increased airplane performance. Thus, they suggested, there could not be any cumulative effect.

Both Muthukrishna et al. (2014) and Kempe and Mesoudi (2014) separately conducted experiments focusing on how group size affects the performance level through larger groups having more models to draw upon. Both of these experimenters examined the performance level obtained after doing a task over several generations in which the group members for the current generation had access to the results obtained by the group doing the experiment task in the previous generation. In both experiments, two conditions were considered: (1) group size $n = 1$ and (2) group size greater than 1, with $n = 5$ for the experiments by Muthukrishna et al. (2014) and $n = 3$ for the experiments by Kempe and Mesoudi (2014).

Both experimenters found that the performance level by a group of size $n > 1$, after several generations of doing the task that was part of the experiment design, was greater than what occurred with a group of size $n = 1$ for each generation. Both sets of experimenters interpreted their results as providing support for the treadmill model, because the groups of size $n > 1$ performed better than the groups of size $n = 1$.

A notable problem with these experiments is that it is unclear whether they really test the treadmill hypothesis. They test its main prediction, but even if the prediction is verified, the question remains as to whether it is caused by the proposed treadmill process or by some other process.

The experiment that is closest in form to the process encoded in the formal treadmill model, and thereby the most in-

interesting in this context, is the one done by Derex et al. (2013). In their experiment, participants were given the task of drawing either a simple (arrowhead) or a complex (fishing net) artifact. The simple task provided a lower but more certain score than the complex task. In each round, participants had access to the outcome of the previous generation (up to the third round; they also had access to the initial state of the “cultural package”) when choosing and drawing either of the two artifacts and received a score on the result. Interestingly, this means that both skillfulness and complexity are represented. Groups of 2, 4, 8, or 16 persons were used to determine the effect of group size on performance. They concluded that the results generally support the treadmill model, because larger groups were more likely to maintain the complex artifact and to maintain both artifacts simultaneously.

The experiment has been critiqued by Andersson and Read (2014) on two grounds. First, the experimental data show that group size actually affects the performance of the groups negatively. Second, a simple sampling effect, without any social interaction involved, would predict a stronger dependence on group size than what was actually observed. Derex et al. (2014) responded to this critique by stating the following: “Our initial analysis showed that the probability of maintaining the complex trait within a group is positively affected by group size. Even if explained by sample size effect, this supports the group size hypothesis: sample size effect is expected to be the main mechanism by which group size affects cultural evolution.” But is it really significant to simply show experimentally that larger groups are more likely to retain a skill? It hardly seems surprising that we should see such an effect. Small groups are more sensitive to noise—for example, the risk of accidentally losing a skill—which is not what the treadmill effect is about.

It is notable that, in terms of absolute performance, increasing group size from 8 to 16 actually produced deterioration in performance; see figure 3 in Derex et al. (2013), figure 1 in Derex et al. (2014), and figure 1 in Andersson and Read (2014). Does this mean that performance decreases, or at least levels off, for group sizes on the order of 10? One may certainly ask how that could be the case. One reason might indeed be that the unrealistic separation between skillfulness and skill complexity in Henrich’s model (see “Treadmill Model”) cannot reasonably be repeated in an experiment. Although this is not explicitly studied by Derex and colleagues, the way that scores are calculated suggests that an artifact with a higher score (higher z) must also be considered to be more complex (higher α) than an artifact with a lower score. That may be reasonable, but it does not correspond to Henrich’s model, and this underscores an important limitation. In the experiment, when the score (skillfulness) increases,⁸ complexity also increases, leading dynamically to what corresponds to a decrease in $\Delta\bar{z}$.

8. Or, more correctly, the rate of skillfulness increase.

Testing the Premises of the Treadmill Model

Validating predictions is necessary but not sufficient for ensuring the soundness of a model. There are many ways to generate any given macroscopic pattern, so the only way to tell whether we found “the right way” is to look at the premises. But this is not just about validation. It is just as much a matter of determining what the model is really about. Uncovering the premises of the treadmill model is, however, far from straightforward. The formal component of the model is as opaque as it is elegant, while the just-as-important informal component has largely been ignored. In this section, we will cover some explicit tests of the premises of the formal part of the treadmill model, as well as some of its elaborations and extensions.

We begin with Vaesen’s (2012a) test of two salient and strong assumptions of the treadmill model, citing their identification by O’Brien and Bentley (2011), but also by Read (2006): (i) the use of a Gumbel distribution, both to model the maximum skillfulness in the population and the skillfulness that results from imitation of a role model, and (ii) the way that role models are selected in skill imitation.

A central premise of the treadmill model is the assumption of a very strong skill bias in role model selection; see “Treadmill Model.” This assumption is embodied in the use of a Gumbel distribution for generating the skillfulness that will be imitated in the next round of imitations (see fig. 2), which greatly simplifies the mathematical derivations and is critical for making it analytically tractable. But what happens if we relax this assumption? It seems unreasonable that imitators can actually identify the best role model across populations possibly in the thousands. Besides, even if they could, the further implication is that this best role model would then have to train the entire new generation. Because the assumption is clearly unrealistic, it is important to test for robustness of the model to the use of other distributions.

Henrich (2004) clearly realizes the need for testing this assumption for robustness, and he does so by using a standard logistic equation in place of the Gumbel distribution. He finds that this did not qualitatively affect the behavior of the model. He also states that, considering the fact that he is investigating losses of skillfulness, the use of a Gumbel distribution errs in a direction that works against his hypothesis: the Gumbel distribution is biased for highly skillful role models, and its use should make cumulative evolution more likely (Henrich 2006). Vaesen (2012a) further tests the assumption of a Gumbel distribution by using a Normal distribution of role model skillfulness, which would certainly be a more standard assumption. He finds that the qualitative behavior does remain intact, although there is a lower risk of cultural loss, which is in line with Henrich’s claim that his assumption was legitimate, because it worked against his hypothesis.

But due to how Henrich’s model is formulated, the Gumbel distribution also becomes a model of the outcome of im-

itation events, a usage for which it entirely lacks a motivation. Although this dual use of the Gumbel distribution is not immediately obvious in Henrich's (2004) model, these two roles become clearly separated in the agent-based models introduced by Powell, Shennan, and Thomas (2009, 2010). Role model selection is here simulated explicitly and involves both a vertical and an oblique element. In the latter mode, role models are selected among agents in the population with a higher z -value, with a probability that is proportional to the difference in z -values. The selected role model's skill (represented as its skillfulness) is then transmitted using Henrich's original model, which employs the Gumbel distribution. But since role model selection is now explicit, Henrich's original motivation for using the Gumbel distribution drops out. At the same time, the choice of a Gumbel distribution and the use of a strong skill bias in oblique transmission also becomes more problematic, because the aim here is to account for increases in cultural complexity. O'Brien and Bentley (2011) point out that this choice, now without motivation, works for the hypothesis and that a more conservative choice of transmission would be advisable. Vaesen (2012a) tests the effects of relaxing the assumption of strong selectivity by using conformist and random copying rather than skill-biased transmission and finds that their results do depend in an important way on a strong skill bias in transmission.

A point of contention that has already been mentioned concerns the prediction of the treadmill model of an indefinite increase in skillfulness in the adaptive regime: the model covers only one round of imitation, and if skillfulness increases ($\Delta\bar{z} > 0$), a simplistic linear extrapolation leaves us with indefinite increase. Even if we eliminate the common misunderstanding that this predicts an indefinite increase in complexity (see "Treadmill Model"), this is still clearly unrealistic (the Price equation is not a model of evolution; it is a model of change over a single generation). There is no reason to believe that Henrich tried to be realistic here, because increases in both skillfulness and complexity were outside of the scope of his study. Furthermore, nothing indicates that Henrich's linear extrapolation was intended as a more general model of evolution, and it is quite clear that, if we want to generalize the model, we would need to formulate more sophisticated models of the evolutionary dynamics.

We have already discussed how Mesoudi's (2011a) critique of indefinite increases in z is negatively affected by his conceptual confusion between skillfulness z and complexity α (see "Treadmill Model"). If we interpret Mesoudi's model as an attempt to generalize the model to deal with increases in z , we see that he does this by attaching a cost to z . But while it is reasonable to attach a cost to complexity (more components, interactions, and discriminations to learn; more material, maintenance, and so on), a high degree of skillfulness would, to the contrary, more likely lead to lower cost (e.g., higher efficiency, fewer failures, and higher usefulness and durability). This, by necessity, evades Mesoudi, because the muddled dis-

inction between complexity and skillfulness leads him to cite examples of increasing α as examples of increasing z .

So cost may not be the most relevant way of dynamically bounding z . It appears that it would be more relevant to internalize (i) the returns to increasing skillfulness and (ii) the simple fact that it generally becomes harder and harder to increase skillfulness further. A limit example of the former would be trivial skills, such as berry or turtle picking. It is easy to pick berries and turtles, and the supply is limited by external factors, so increasing skillfulness only pays off so far. An example of the latter could be running a 4-minute mile; some skills are so strongly based in motor capabilities that increasing skillfulness rapidly becomes very hard.

Read (2006) criticizes Henrich's (2004) choice to graph the implications of his model for values of $\alpha/\beta > 4$, stating that this leads to estimates of required population sizes that are vastly overblown. To illustrate how this leads to unrealistic assumptions about empirical premises, Read uses percentiles, assuming a normal distribution for degrees of skillfulness, stating that, if we view the top 1% of a population as "highly skilled," then we would expect to find five individuals in this category in a population of 500. If we view the top 5% as being very skillful, then we would expect to find at least one individual in this category in a group of 20 individuals, which is in the range of a normal forager residency unit. This points us to another contentious aspect of the treadmill model: what does population size really correspond to?

Recent contributions have made elaborations to the treadmill model by relaxing assumptions about population structure. Kobayashi and Aoki (2012) relax the implicit assumption of Henrich (2004) of using discrete non-overlapping generations, and they investigate theoretical implications of subdividing the population into demes of more strongly interacting networks of acquaintances. They find that their model predicts substantially smaller required populations than Henrich's (2004) model, which they attribute to the use of overlapping generations. They also find that, if oblique transmission is limited and structured by a heterogeneous network structure—that, for example, can be argued to represent a subdivision of a population into a metapopulation of residential groups over a geographical area—then the gross population size N produces a weaker effect. In the model that they use, parameters for degree of interconnectivity and rate of innovation become more important.

Baldini (2013) also elaborates on population structure and comes to similar conclusions. He finds that the effects of the size of networks within the population and rates of error in transmission overshadow the effects of population in his model. The only strong effect of population size that he detects enters when innovations are rare and transmission is easy, so that the size of the whole population becomes limiting for the availability of improvements to accumulate. Nakahashi (2014) follows the line of investigation of Kobayashi and Aoki (2012) but also relaxes the assumption of strong skill bias by testing, for example, for vertical transmission and conformist bias.

Powell, Shennan, and Thomas (2009, 2010) place the treadmill model in a geographical setting, and they explicitly split the population into interacting subpopulations. They thereby elaborate on the notion of population size as a measure of the population of interacting social learners. Recognizing the unrealistic possibility of considering populations of thousands of foragers within which role model selection can effectively be made (Read 2006), they implement a model of migration across space that keeps role model selection local while bringing in nonlocal population effects in a more realistic manner. Beyond basically seconding Henrich's (2004) conclusion that population size makes or breaks cultural accumulation, Powell, Shennan, and Thomas (2010:157) add that "the level of cultural skill that can be maintained in subpopulations is related to the density/migratory activity of those subpopulations."

Complexity has so far been quantified empirically using Oswalt's (1976) techno-units (e.g., Collard, Buchanan, and O'Brien 2013; Collard, Kemery, and Banks 2005; Collard et al. 2013a; Kline and Boyd 2010; Read 2006, 2008). Querbes, Vaesen, and Houkes (2014) argue, with the aid of a model, that techno-units are a poor proxy of complexity (in its interpretation as complicatedness), showing that the behavior of models changes in important ways if the number of interactions between parts, rather than the number of parts, is considered. They also come to the same conclusion as Andersson (2011, 2013), that the degree of interconnectedness between parts greatly affects sensitivity to transmission errors.

It is also easy to see how any measure of cultural complexity performed on the physical form of finished artifacts could be systematically misleading. Not all complexity associated with the cultural transmission of an artifact is necessarily (or even typically) expressed as complexity in its morphology, nor do all skills even involve a distinct physical outcome. Consider, for instance, the lithic beads produced at Khambhat, India, as described by Bril, Roux, and Dietrich (2005). These spherical beads are exceptionally simple from the point of view of techno-units, but from the point of view of the technique used to produce them and the information that needs to be culturally maintained, they are exceptionally complex.

Explanatory Role of the Treadmill Model

To offer a more detailed analysis, we now differentiate between two ways in which the treadmill hypothesis can be interpreted: a strong and a weak interpretation. The strong interpretation is that Henrich's hypothesis regarding population size is the main driver of cultural complexity. The weak interpretation is that it represents one causal factor among many.

The strong interpretation relies critically on a clear signal in data across forager groups. Reviewing the empirical picture, which has been substantially augmented over the past decade, we find that the strong interpretation of Henrich's (2004) hypothesis can be rejected. Negative evidence reasonably weighs more heavily if the question is whether Henrich's

hypothesis is the dominant determinant of cultural complexity. Even ambiguity in data is devastating to the strong interpretation, and the only conceivable way to salvage the model in a strong interpretation appears to be the possibility that a signal would emerge from data where other, and arguably more accurate, concepts of complexity and population are used. We deem this somewhat unlikely on the grounds that there is substantial evidence of correlation between (techno-unit) complexity and other group features, most importantly a proxy of environmental risk (see "Testing the Predictions of the Treadmill Model").

The weak interpretation does not correspond to any single model formulation. It concerns the wider question of how different causal factors would fit together to produce explanations. Is, for example, the treadmill model valid in certain types of situations? Is it valid only for certain types of skills? Does it represent one constraint on complexity among many, such that in any given situation, some other constraint could modify or preempt it? We argue that not only the treadmill model but also the other causal explanations that have been proposed should be vetted in weak versions. The premises of models become particularly important as the question is no longer just whether the model is valid but also what it depicts and how it can fit together with other models.

An important reason for why the strong interpretation enjoys so much momentum today appears to be a widespread and tacit assumption that the treadmill model represents the only conceivable process that could link population size to complexity. Authors have tended to get away with claims of having supported Henrich's result as soon as they have demonstrated an increase in measures of skillfulness or complexity with increasing population sizes. This is true not least for recent experimental work; see "Testing the Predictions of the Treadmill Model." Failure to properly consider the premises upon which models and experiments are constructed exposes us to the risk of both false corroborations and falsifications.

So what is wrong with the treadmill model? We believe there are many things right, but that, on the background of viewing the treadmill model in a weak interpretation of its explanatory role, two things are missing so far.

First, we need a clear picture of what the treadmill model does—and does not—predict and assume. We hope to have contributed to clarifying this picture somewhat, although it is clear that more work is needed. The treadmill model is elegant, but it is also highly opaque. In particular, we have pointed out that the subdivision between skillfulness z and complexity α and the lack of an explicitly formulated view and model of how α really evolves have been the source of serious confusion. This should be addressed, such as with more explicitly formulated simulation models that map more directly to anthropological and archaeological data at the level of premises: simple but more transparent models, affording sufficiently resolved representations of interactions, mechanisms, and cultural organization, not least with regard

to demographic structure. In the next section, we will confront some of the premises reviewed and uncovered here in more empirical detail to see how this may provide a clearer view of the treadmill model in its immediate context.

Second, we need a careful analysis of how the treadmill model would interact with models based on other factors. We will outline such an analysis of the wider context of the treadmill model in the last section.

Treadmill Model in Its Immediate Context

Let us begin by noting that Henrich's (2004) model fundamentally assumes that societies will have cultural traits whose complexity α is near the limit of what they manage to maintain (see also "Treadmill Model"). Consider a situation where this would not be the case: say we have a population of 20,000, where all skills are sufficiently easy to imitate that they would need only a population of 20 to ensure their stability. If we cut this population down to 5,000, or if we double it to 40,000, the treadmill model would then obviously predict no effect. We contend that this is exactly the case in forager societies.

Why do we think this is the case? There are two reasons, each sufficient on its own. The first reason is that forager groups are largely self-sufficient units, and the whole battery of skills that they rely upon must be robustly present in each group. The populations of such groups range in the tens rather than the hundreds or thousands. Foragers simply do not maintain skills that only a few individuals within a large network of groups can master; they maintain skills that just about anybody that is properly enculturated will master. This may, of course, be different in producing rather than procuring societies with developed craft specialization, but that does not relate to questions about human evolution in the deep past.

The second reason is that, although forager groups, at least sometimes, interact strongly (Hill et al. 2014), the question is whether and, if so, to what extent they interact in the right way to maintain skills collectively between such groups. To transmit sophisticated skills, learners must engage intimately with role models. One way of expressing the reasons for this is that such skills demand what Tostevin (2012) refers to as an emic experience of the practice of performing the skill. Emic, as opposed to etic, aspects of skills are part of the personal experience and cannot be explicitly instructed, so the only way to transmit them is by close and repeated interaction over long periods of time.

The move that this calls for is a subdivision of Henrich's grand N parameter into communicating subpopulations of a much smaller size M in metapopulations. This has been implemented so far in two main ways: (i) by the introduction of a network structure (Baldini 2013; Kobayashi and Aoki 2012) and (ii) by the introduction of migration between subpopulations (Powell, Shennan, and Thomas 2009, 2010).

The former implementation does not take the constraint on emic aspects of skill transmission into account and should

be expected to exaggerate the impact of N ; careful role model selection and transmission of emic knowledge, although constrained, still takes place across the entire population pool of size N . It is not sufficient to introduce only a network structure; one must also introduce heterogeneity in the types of information that can be transmitted between different parts of the network.

The latter implementation, by Powell, Shennan, and Thomas (2009, 2010), does incorporate this constraint; at least it does so potentially, depending on how one interprets what it means to migrate. Transmission and role model selection here—more realistically, in our opinion—take place only within each subpopulation of size M . The model of Powell, Shennan, and Thomas (2009, 2010) shows results that are in accordance with our synthetic assessment here: what corresponds to N (which relates to the density of subpopulations) becomes unimportant beyond a certain low value. Below this point, N is important, but it is unclear exactly why it is important.⁹

For this reason, we find it likely that the treadmill model acts to constrain complexity based primarily on sizes M of residential groups, not the sizes N of larger networks of groups (although there is likely some interaction between these two quantities). The treadmill mechanism may thereby impose a bound on the size of groups downwards, and it may well be a reason why larger residential units over time were able to maintain more complex cultural systems. In this form, the model may apply to why we see craft specialization only in societies with large populations: craft specialization fragments the population into smaller "skill populations," each of which needs to be of sufficient size.

Treadmill Model in Its Wider Context

We will now briefly review several models representing other hypothetical drivers and constraints on cultural complexity and then reflect on how they might be integrated. A largely parallel lineage of models invokes the concept of a cultural "ratchet" (e.g., Pradahan, Tennie, and van Schaik 2012; Tennie, Call, and Tomasello 2009; Tomasello 1999) rather than a "treadmill." The choice of metaphor is significant, as the focus here is directly on cultural gains rather than losses. Models in this tradition focus to a large extent on the cognitive capabilities that underpin cumulative cultural evolution, and the studies cover the whole spectrum between animal and human culture (e.g., Dean et al. 2012, 2014; Vaesen 2012*b*). References to the treadmill model are frequently seen in this lit-

9. It could be because migration in their model is density dependent, so that N actually impacts the effective subpopulation sizes, which, in turn, is important. The effect of this migration is expressed by Powell, Shennan, and Thomas (2010) as analogous to increasing the sizes of the subpopulations, because both have the effect of increasing their internal variance in skillfulness.

erature, but mostly as a way to provide a wider context that is not explicitly integrated.

One constraint on cultural complexity that has recently come to the fore in this “ratchet tradition” is the role of fidelity in cultural transmission as a constraint on cumulativity (e.g., Tennie, Call, and Tomasello 2009; Tomasello 1991; Whiten et al. 2009a, 2009b). Originally, the role of fidelity was mostly discussed in loose terms, where transmission was said to have to be “sufficiently faithful” for cumulative evolution to be possible (e.g., Claidiere and Sperber 2010). But more recently, a number of studies have investigated various aspects of fidelity in more detail. For example, Enquist et al. (2010) find that fidelity and the number of role models strongly affects the longevity of traditions, Lewis and Laland (2012) find that high fidelity is crucial for cultural cumulativity, and Andersson (2011, 2013) finds that fidelity imposes a cap (a “glass ceiling”) on the amount of information that can be carried by cultural transmission.

The basic treadmill idea (i.e., that adaptation must constantly work against a pressure of transmission loss) is indeed far from new in evolutionary thinking, even if it is less salient in biology than in archaeology and anthropology. For higher life forms, genetic transmission is anything but noisy (e.g., Drake 1998), so it makes sense to view the evolution of higher life forms from the point of view of sorting variants without worrying about whether they will be lost on a large scale despite providing adaptive advantages. This is not the case, however, if we look at primitive and simple forms of biological and prebiotic evolution (e.g., viruses and RNA replicators), for which Eigen and Schuster’s “quasispecies model” (Eigen 1971; Eigen and Schuster 1977; see Domingo et al. 2012 for a recent review) depicts precisely a “treadmill” of loss of genetic information.

Andersson (2011, 2013) adapts this quasispecies model to cultural evolution to argue for a treadmill connection between transmission fidelity and cultural complexity. The argument by Andersson is that cultural evolution must be viewed as primitive in this sense: mechanisms for high fidelity—cultural or physiological—must have emerged over time, and we cannot assume that they were there to begin with. Cultural evolution must initially have been constrained by low fidelity in a way that is similar to what is the case for biological systems with low fidelity transmission. What corresponds to skills here have a size, corresponding to their complexity and how large a target for error they are. This corresponds closely to the α parameter and the notion of complexity in the treadmill model. Andersson and Törnberg (2016) combine the fidelity argument embodied in the models by Andersson (2011, 2013) with the treadmill model (Henrich 2004) to yield a synthetic model where imitation fidelity and population size (via skill bias and variance) act together as evolutionary constraints. The results indicate that the latter is important for very small population sizes ($N < 15$) but that its impact tapers off rapidly for larger population sizes.

Fidelity, of course, is challenging to measure in the field, and so it has mostly been done in laboratory settings, such as

on chimpanzees and children (e.g., Horner et al. 2006) and fish (Laland, Atton, and Webster 2011). As argued in different ways by Andersson (2011, 2013), Sterelny (2011), and Tostevin (2012), fidelity in cultural transmission is, however, not just expressed in microlevel contacts between individuals. Across generations, more macroscopic processes become involved—pedagogy, learning environments, the use of generative logic (such as in language and cultural idea systems, like kinship systems; see Leaf and Read 2012a), and so on—and the modified environment itself is implicated in the process. It is also likely that hominin evolution has brought about new and unique adaptations in the domain of pedagogy. Csibra and Gergely (2009, 2011) propose an emerging package of such adaptations—a hominin “natural pedagogy”—that goes beyond what we may simply characterize as imitation or emulation, or what we can just quantify as fidelity. All these factors would also affect the notion of imitation quality used in the treadmill model.

The innovation cascade model introduced by Schiffer (2005), which recalls similar models in the social sciences (e.g., “exaptive bootstrapping”; Andersson, Törnberg, and Törnberg 2014; Lane 2011; Lane et al. 2009) adds another aspect to the evolution of complexity, depicting an intrinsic drive for change, which includes increasing complexity as a possibility that stems from how culturally maintained systems are organized and used. Key to this type of process is that, as the system (e.g., a technology) is used, new uses, problems, and shortcomings will be revealed in the process, leading to further modifications and reattributions of functionality (i.e., exaptation; see Gould and Lewontin 1979).

What we initially referred to as the “traditional connection” between population size and cultural complexity has also been pursued further recently. Stiner, Munro, and Surovell (2000) and Stiner and Kuhn (2006) combine Flannery’s (1969) broad-spectrum revolution with foraging theory (Stephens and Krebs 1986), extending the explanation that Flannery offered by using it as a general framework for understanding not only why cultural complexity should increase but also the directions in which it might be expected to go.

Drivers have also been proposed in relationship to adaptive responses to risk in resource procurement and tool complexity (e.g., Collard, Kemery, and Banks 2005; Oswalt 1976; Read 2006; Torrence 1983, 1989, 2000) and, to a lesser extent, settlement mobility (Shott 1986). Read (2008) has shown that about 95% of the variability in the complexity of hunter-gatherer tools may be accounted for by an interaction model incorporating risk and mobility and taking into consideration the two different resource-procurement strategies employed by hunter-gatherer groups; thus, hunter-gatherer tool complexity may be viewed as an adaptive response to environmental conditions.¹⁰

10. Collard et al. (2011) have argued recently that the apparent adaptive response to risk may be a function of the magnitude of the difference in risk experienced among hunter-gatherer groups. They found that, for hunter-

It is notable that these results for hunter-gatherer groups do not extend unmodified to herders and small-scale farmers (Collard et al. 2012, 2013*b*). Apparently, the dynamics affecting tool complexity among food procurement groups such as hunter-gatherers are not the same as for tool complexity among food production groups. One possible difference could be that the effect of population pressure could be more telling among the latter if tool complexity correlates primarily with rates of food production, which correlates with population size, whereas risk reduction has been found to correlate with tool complexity among the former. Risk is an exogenous factor for hunter-gatherer groups and hence does not correlate with population size. Another possible difference is that more sophisticated institutions in food-producing groups may mitigate the effects of small population size on the transmission of culture (see also Andersson [2013] and Henrich [2009] for a similar argument about institutions and the effects of low-fidelity transmission). A third difference could be that the presence of craft specialization and developed trade in food-producing societies makes for a fundamentally different innovation and cultural transmission process.

Yet even this may not be sufficient. Culture transmission involves a process by which each new group member is enculturated into an ongoing cultural system. We do not “choose” whether or not to become enculturated or even necessarily by whom (but see Henrich and Broesch 2011). Culture involves idea systems (Leaf and Read 2012*a*), not the artifacts that are the instantiation of cultural systems. As the archaeologist Irving Rouse (1939:15) puts it: “Culture does not consist of artifacts. The latter are merely the results of culturally conditioned behavior performed by the artisans.” Cultural evolution involves change in idea systems, such as the kinship terminologies that are part of a group’s kinship system. Such a terminology has an underlying generative logic that accounts for its structure and organization (Leaf and Read 2012*a*; Read 2001, 2007), so change occurs through its generative logic (see, for example, Read 2013). In brief, models of cultural evolution need to consider change in idea systems, more in analogy with the evolutionary developmental conception of variation as variation in developmental trajectories (e.g., Arthur 2011) than with an atomistic populational view. This will require models of a different sort than those used to model the evolution of biological traits (see Lane et al. 2009; Read 2010, 2012*a*).

gatherer groups in the Northwest Plateau with similar risk, the correlation between risk and tool complexity and diversity was attenuated. However, as is well known, reducing the variability in an explanatory variable will necessarily attenuate the correlation between that explanatory variable and any measure of concern in a regression model. Logically, if there were no risk difference among the groups making up the data set, then any variability in a response variable across the groups in the data set would be due to factors other than risk. Thus, the effect they observe does not relate to risk and tool complexity but is simply a necessary property of linear regression models.

Conclusions

What emerges from this overview and from observations made earlier in the text is that the main proposed drivers and constraints on cultural complexity made to date are as follows:

1. Population as it pertains to social interaction; this is the “treadmill version.”
2. Population pressure (i.e., the cultural-complexity-mediated interaction between populations and resources); this is the “traditional version” holding that higher cultural complexity allows, or is needed, for larger populations.
3. Adaptivity; cultural complexity as an adaptive response to environmental variables and cultural organization.
4. Transmission fidelity and its nonlinear dynamical effects on evolution.
5. Cognitive preconditions for imitation, emulation, and pedagogy and, thereby, human cultural cumulativity.
6. Pedagogy and similar mechanisms as cultural institutions.
7. Bootstrapping processes that drive the emergence of complex organization.
8. Generative logic in systems of concepts serving to organize innovation and simplify error correction in transmission.

It is not obvious how to combine these different factors, and this is very likely because of a tendency to pursue explanations individually in strong interpretations. If several of these factors were important together, as we think is the case, such a debate would resemble an argument about whether the wings, the fuselage, or the engine is what really makes an aircraft fly. We think this is precisely what we see today. All of these propositions are well argued, and they find at least some support at the level of premises or predictions, but none seems to be capable of persistently convincing a majority.

This carries important implications for determining how to proceed scientifically. While in many cases it is a scientifically sound practice to isolate and explore causal factors, there is a big difference between expecting one such factor to be “the winner” in the end (a strong interpretation), and expecting it to be part of an explanation where the factor in question interacts with other factors (a weak interpretation).

If we go with a weak interpretation of the explanatory role of models, we must also realize that the collective action of a system of interacting factors is emergent. That is, it cannot be expected to be understandable as a simple juxtaposition of individual factors.

For this reason, we think that more attention should be paid to interactions among the factors that have been proposed and that we have listed above. This approach contrasts with a competitive mindset (which is arguably prevalent today), where favored explanatory hypotheses are championed to the exclusion of others (or, more accurately, where other factors are dwarfed, subsumed, or even ignored by the preferred explanation).

On the most aggregated level, it appears highly likely that the course of cultural evolution relates to a combination of endogenous and exogenous processes. The natural envelope of such an aggregated model would be along the lines of “the traditional connection” with interaction between culture and the environment mediated through population size. This interaction between external and internal processes, however, should not be construed as an opposition between two clearly delineable levels of organization.

We should also not lose sight of the “vertical,” hierarchical organization of the cultural domain. One notable aspect of the factors that we list above is that they vary widely across scales of time and levels of organization and, in many cases, also differ in whether they take a top-down or a bottom-up perspective on culture.

Anthropology and archaeology are far from alone in experiencing strong tension between models formulated on different levels of organization and time scales and between models taking bottom-up and top-down approaches. Lessons can be taken from “developmental evolutionary theory” (see Laubichler and Maienschein 2013), such as niche-construction theory (e.g., Odling-Smee et al. 2003), generative entrenchment (Wimsatt 2013), new approaches to transitions in natural history (Erwin and Valentine 2013), and, critically, innovation in the process of innovation that characterizes hominin cultural evolution leading to *Homo sapiens* (Read 2012a; Read, Lane, and van der Leeuw 2009). Recent sociological theory of innovation and technical change (Geels 2002; Geels and Schot 2007; Lane 2011; Lane and Maxfield 2005) can be integrated in such a view (see Andersson, Törnberg, and Törnberg 2014). From this perspective, most of the drivers and constraints, including those represented by the treadmill model, the fidelity-based models, and the proposed roles of pedagogy, will be increasingly intermixed as more abstract determinants on what is likely, possible, and impossible are developed.

Formal models, such as the treadmill model, are useful but must be kept in perspective. The interpretations made of them must be consistent with their underlying assumptions, and their assumptions must be shown to be consistent with relevant, empirical data. We argue that the evolution of cultural complexity is a multifaceted process and does not lend itself to single-factor explanations like the treadmill model. Progress in our understanding of the development of cultural complexity requires that we better understand what we mean by cultural complexity and what actually evolves when we refer to the evolution of cultural complexity, and, perhaps most important, requires that we take an integrative approach to modeling cultural complexity that recognizes its multifaceted character.

Although we have focused on the treadmill model, this is only because of its dominance in the literature today. The treadmill model, as it stands, is too narrow in its scope and is limited by the strong assumptions it makes. It does not provide a universal explanation for the variability and evolution of cultural complexity, hence our rejection of the strong in-

terpretation of the treadmill model. We advocate, instead, what we refer to as a weak interpretation of its explanatory role (see “Explanatory Role of the Treadmill Model”), not only for this model, but for any proposed uncausal explanation of cultural complexity; Andersson and Törnberg (2016) represents an effort in this direction. Cultural complexity points, not to any single factor, but to the consequences of a system of causal factors.

Comments

Kimmo Eriksson

Mälardalen University, School of Education, Culture and Communication, Box 883, SE-72123, Västerås, Sweden (kimmo.eriksson@mdh.se). 18 VIII 15

At the end of their thoughtful target article, Andersson and Read conclude that formal models of cultural evolution are “useful but must be kept in perspective.” As a mathematician with a great interest in social science, I have some experience of working with such models. Based on this experience, I very much agree with the “but” part of the above conclusion. I see a clear tendency in the cultural evolution literature to put too much trust in the value of formal models. Specifically, it seems to me that researchers often spend too little effort in analyzing the real-world processes by answering questions such as, “At a concrete level, what are the cultural traits I am interested in? In what respects do they vary? How does variation seem to arise? How do these traits spread?” In my experience, attempting such a concrete analysis makes one realize how complex and multifaceted these processes are and how little we really know about them—a humbling insight. But it seems to me that modelers tend to pay little attention to this step of the research process and quickly jump to setting up an abstract and mathematically convenient model. Such jumps come with an increased risk of the model creating confusion when applied to real-world cultural evolution.

The treadmill model discussed by Andersson and Read is indeed problematic in this way. The most fundamental assumption in this model is that any particular cultural trait can be expressed at different levels of skillfulness measurable on a unidimensional scale. Within the context of an abstract mathematical model, this assumption seems intuitive and innocuous. (After all, there is surely some variation between different instances of what people would recognize as the “same” trait; what is the harm in referring to the dimension of variation as “skillfulness”?) But intuitiveness in the abstract is not good enough, because the model is, in the end, meant to explain some pattern of concrete events in cultural history. For the concrete cultural traits that are relevant in the context to which the model supposedly applies, we need an explicit interpre-

tation of what variation in skillfulness looks like. Andersson and Read do a good job at explaining the difficulty of nailing down such explicit interpretations. This is bad news for the model, because the concepts of cultural traits and the skillfulness with which they are expressed are its fundamentals. Only when these concepts are made sufficiently concrete are we in a position to evaluate the validity of the other main assumptions of the model: that cultural traits spread through imperfect imitation; that the currently most skillful version of the trait in the population is the one that is imitated; that the skillfulness of imitated traits varies but, on average, decreases compared with the original; and that the average decrease in skillfulness is a reasonable way of defining the “complexity” of a trait.

Each of these assumptions can and should be questioned. In my view, questioning assumptions is a very useful endeavor. I think the future of mathematical modeling in cultural evolution research lies in a shift of emphasis. So far, the main use of mathematical models has been to investigate the dynamics of cultural evolution by analyzing the dynamics that arise in the models. Most of these investigations I would consider to be premature, because they are based on assumptions and conceptualizations of questionable validity. Instead, I would like to emphasize the use of formal modeling as a way to develop the links between empirical knowledge and theoretical conceptualizations. For instance, the primary use of the treadmill model would then be to serve as the basis for a scholarly discussion of how to conceive of cultural traits and how they spread and change. Note that such discussions would naturally also involve empirical researchers who may be daunted by the mathematics involved in the current focus on dynamical analysis. A shift in focus from the mathematical analysis to the underlying assumptions should lead to a more engaging and productive discussion.

I am hopeful that such a shift may ultimately lead to better models and better understanding of their scope of application. Andersson and Read’s article is a very good move in this direction.

Ian Gilligan

Sydney, Australia (ian.g@bigpond.net.au). 30 VIII 15

The treadmill model attempts to link social evolution to population size on the basis of the fact that cultural transmission of skills requires imitation that is usually imperfect. It proposes that a minimum population size is needed to ensure sufficient innovation to compensate for a constant drain due to errors in transmitting knowledge and skills. When group size becomes too small, the rate of loss will outstrip replication and innovation. The result is flawed transmission and failure to outrun the treadmill. This can lead to maladaptive losses and depletion of technologies, compromising a society’s evolutionary prospects.

The seminal article (Henrich 2004) showed how the treadmill model might explain a major “puzzle” in anthropology: the apparent social devolution of Tasmanian Aborigines during the Holocene. The problem is that this retrograde development—or simplification—should not happen within an evolutionary scenario. If it does happen, and yet a society still manages to survive—or worse, thrive—then it potentially questions the validity of social evolution.

Andersson and Read subject the treadmill model to a welcome critique and show that, despite its popular appeal, it can be refuted by relevant data (at least in its “strong” formulation). They are inclined nonetheless to advocate for its continued retention in a “weak” version. I will return later to their assessment, but I first wish to highlight a substantive failure of the model in Tasmania that, in my view (and for a number of reasons), warrants a more severe rejection.

Foremost among the “maladaptive” losses are the disappearance of bone tools from the archaeological record and the ethnographic absence of “cold weather” clothing. The two may be connected, with bone tools serving as awls for sewing animal skins into warm garments. The anomaly is attributed to Henrich’s treadmill effect: when Tasmania was severed from mainland Australia by rising sea levels at the end of the Pleistocene, the effective population size was reduced below his minimum. While distancing himself from archaeologist Rhys Jones’ infamous conclusion that the inhabitants were “doomed” by their isolation (Jones 1977), Henrich regards Tasmania as both a key illustration and a crucial test of his model. I agree.

Andersson and Read mention my work on Tasmanian clothing, but a few issues need explicit discussion. First, Henrich’s claim that their clothing was insufficient (“maladaptive”) is an absurd proposition, consistent with his general neglect of biology and climate. With hypothermia, survival times are measured in hours, not millennia (Gilligan 2010). He ignores the role of differing biological adaptations between populations that will impact on cold tolerance and hence on clothing requirements. Such differences are well-documented in physical anthropology in populations including Australian Aborigines and others he cites, such as Fuegians and Andaman Islanders (e.g., Gilligan and Bulbeck 2007; Gilligan, Chandraphak, and Mahakkanukrauh 2013). These considerations are crucial for understanding why Tasmanians in the Holocene managed perfectly well despite wearing less clothing than their neighbors (Gilligan 2007*b*).

The Pleistocene situation is especially interesting because, as Henrich (2006) concedes, his model collapses if Tasmanians were not enmeshed within wider social networks to the north. Yet we have no evidence for ongoing connections. On the contrary, given local conditions on the exposed land bridge as the glacial maximum approached (Colhoun 2000; Hope 1978), the founding population may have remained culturally isolated since their arrival in the region. Likewise, Henrich’s supposition that Tasmanian tribes were not culturally differentiated in the Holocene is difficult to sustain

in view of linguistic evidence for 12 Tasmanian languages (Bowern 2012).

Moreover, dramatic innovations in Tasmania during the Pleistocene challenge the treadmill model. Even if the Tasmanians were not completely isolated, these technological and other developments began earlier and were more comprehensive than those witnessed elsewhere on the continent. This pattern, together with the Holocene disappearance of bone tools, is explicable without invoking demographic factors (Gilligan 2014).

Henrich's ethnographic contrasts with Fuegians and Andaman Islanders are revealing. He attributes the Fuegians' more substantial clothes to their alleged wider social connections (Henrich 2004). Yet the simple difference in climate due to higher latitude is sufficient to explain the difference with Tasmania (Gilligan 2007*b*). Paralleling the Tasmanians, the Fuegians nonetheless wore less clothing than their Patagonian neighbors to the north, probably for similar reasons; their cold tolerance astonished Darwin (1839). We can excuse Darwin's ignorance of thermal physiology in 1832 but not Henrich's in 2004. Similarly, he suggests that the Andaman Islanders benefited from hypothetical connections to Asia (Henrich 2006), but little evidence is discernible in historical or genetic records (Cooper 1989; Wang et al. 2011). And despite highlighting the Tasmanian paucity of clothes, he omits to mention that the technological complexity of the Andaman Islanders did not extend to clothing, which was less than that of their Asian neighbors—less even than that of Tasmanians (Cipriani 1966; Colebrook 1807; Mouat 1863; Temple 1901).

I have deeper misgivings. One is value judgement: aside from using the awkward term "maladaptive," Henrich (2004) actually states that "valuable" technologies were abandoned. The Tasmanians might have disagreed. This raises concerns about veiled ethnocentrism and the nature of the discourse concealed within evolutionary approaches to the Other (Descola 2005; Derrida 2002; Sahlins 2008), not to mention the explanatory weakness of the adaptation concept with culture (Ingold 1980) and even biology (Popper 1972). There is irony, too, in privileging social evolution while discounting biological evolution. I am bothered also by the appearance of elegant mathematics alongside inelegant ethnography. Mathematics lends a scientific "garb" (Husserl 1954), as does evolution. I note that, in their abstract, the authors refer more cautiously to the "elaboration" of cultural complexity.

Andersson and Read are generous in wanting to salvage the treadmill model, albeit in a weaker variant. If the strong version is invalid, it is not clear to me why a weaker version would be less so. Neither should it be shielded within multifactorial approaches, which can similarly have the untoward effect of covering the weaknesses of individual factors. More worrying is how the traction it has gained in the wider world may distract from more nuanced approaches that can better accommodate the interactions between biology and culture.

Michael J. O'Brien and R. Alexander Bentley

Department of Anthropology, University of Missouri, 317 Lowry Hall, Columbia, Missouri 65211, USA (obrienm@missouri.edu)/ Department of Comparative Cultural Studies, University of Houston, 3623 Cullen Boulevard, Houston, Texas 77204, USA (rabentley@uh.edu). 31 VIII 15

With their summary of debates over the role population size plays (or does not play) in the evolution of cultural complexity, Andersson and Read provide a detailed road map through the pluses and minuses of competing models that date back at least to the Enlightenment. Hume (1985 [1777]:382), for example, in response to the view of Montesquieu and others that the population of the ancient world was larger than that of the modern world (Engerman 1997), wrote, "wherever there are most happiness and virtue, and wisest institutions, there will be the most people."

In anthropology, modern debates stem from Shennan's (2001) article "Demography and Cultural Innovation," which was followed by three studies that addressed trait loss in Tasmanian toolkits over an 8,000-year period (Henrich 2004, 2006; Read 2006). Henrich (2004) argued that behavioral information—in the Tasmanian case, information on how to produce certain tools (spear throwers, boomerangs, hafted fishing spears, and so on)—can be lost through processes such as imperfect imitation of a skill (Eerkens and Lipo 2005). A population must continually compensate for this "treadmill of cultural loss" (Kline and Boyd 2010)—a Red Queen effect whereby a population runs faster and faster just to remain in the same place. Henrich's premise was that, on average, a larger population means less cultural loss. Conversely, having fewer people leads to higher rates of loss, especially of "tools that are hard to learn to make, and easy to screw up" (Henrich 2006:776).

Powell, Shennan, and Thomas (2009) used Henrich's model to propose that the explosion of cultural evolution in Europe ca. 40,000 BC, traditionally considered the signature origin of biologically modern humans, could reflect a population increase with no necessary changes in human cognition, counter to a popular view (Klein 2002; Mithen 1996). Powell and colleagues added stochastic and geographic elements to Henrich's model to show how chance clusters of local migrating groups could, by exceeding the crucial population threshold, begin to undergo cumulative cultural evolution over generations.

As Andersson and Read point out, the treadmill hypothesis is now treated as established fact in some circles. Within the last several years, however, a number of studies focused on identifying the drivers of the complexity of tool kits of farmers, pastoralists, and hunter-gatherers from various environmental and ecological zones (e.g., Collard, Buchanan, and O'Brien 2013; Collard et al. 2013*a*) have shown that risk and mobility, rather than population size, are the major factors in terms of increasing or decreasing complexity.

Clearly, as Andersson and Read argue, we cannot assume that any demonstrated correlation between population size

and complexity automatically favors the treadmill model. Demography represents one among several causal factors in how population density and structure affect social learning, as recent discussions of innovation in the developing world have shown (Banerjee et al. 2013; Malakoff 2013). Social-network structure is also a crucial factor (Centola and Baronchelli 2015). The fact that innovation increases superlinearly with urban population size (Bettencourt and West 2010), for example, is partly a result of the face-to-face interaction facilitated by urban life (Pan et al. 2013) but also of how those interactions are structured through assortative mixing, organizations, and communication. In this view, one fascinating research question is how modern media and organizations, as they replaced kinship as the prime organizers of cultural transmission, affected the pace and direction of cumulative cultural evolution (Bentley and O'Brien 2015).

Andersson and Read echo two issues we have raised (Bentley and O'Brien 2011; O'Brien and Bentley 2011; see also Read 2006; Vaesen 2012a). One is the effect of assuming a Gumbel (as opposed to a Gaussian, say) distribution as a model for both the maximum skillfulness in a population and the skillfulness of learners who learn from that maximum. Another is exactly how role models are selected from within a population. It may seldom be the case that potential imitators can find the "best" model (Atkisson, O'Brien, and Mesoudi 2012). Alternatively, instead of imitators copying the best model with some error, each individual is copying the average skill level—averaged across individuals in the group—plus some minor learning error that is normally distributed around zero, either positively or negatively. This standard model yields a random walk in terms of the mean skill level for the group, with stochastic change that can go up or down over time (Bentley and O'Brien 2011). In addition, such a model is unpredictable, in that each random walk is unique. Thus copying the majority, where behavior is continually drawn to the status quo, could make cumulative adaptive evolution merely a matter of drift (Hamilton and Buchanan 2009). Vaesen (2012a) provided a more formal mathematical proof that the cultural-loss hypothesis (Henrich 2004) still holds when assumptions about the selectivity of social learning are relaxed but that cultural gain disappears when social learning is less selective, such as through conformist bias.

The bottom line is that demography is never the universal primary driver of cultural complexity. Hence the selectivity of social learning, or the "transparency" of expertise (Bentley et al. 2014), is among the crucial variables to be measured. Population size is part of this variable, but so too are social-network structures, homophily, and the media of communication. Besides, the social network is also the network of ideas themselves—the path-specific potential for complementary technologies or ideas to be recombined into novel ones (Hidalgo and Hausmann 2009).

When Samuel Johnson stated, "by seeing London, I have seen as much of life as the world can show" (Boswell 1848 [1791]:35), he was referring to both people and ideas. To

turn that into an anthropological example, we could ask, "Were Paleolithic cave art traditions maintained through thousands of years through a continuous transmission chain of generations of expert artists and their apprentices, or through the cave art itself, which could have been imitated at intervals of many generations?" This is a challenging research question that invites a detailed analysis of the pathways of cultural transmission beyond the simple correlation between population size and cultural complexity.

Tim Taylor

Department of Prehistoric and Historical Archaeology, University of Vienna, 1190 Wien, Franz-Klein-Gasse 1, Austria (timothy.taylor@univie.ac.at). 19 X 15

Richerson's *Nature*-validated claim that "group size determines cultural complexity" (2013:351) puts one in mind of E. B. Tylor, writing in the preface to the second edition of *The Aborigines of Tasmania*: "That these rude savages remained within the present century representatives of the immensely ancient palaeolithic period, has become an admitted fact" (Tylor in Ling Roth et al. 1899:vii).

The comparison underscores the fact that Andersson and Read's article should not have been as necessary as it clearly is. While welcoming it warmly, one has also to be concerned by the growing popularity of simplistic and reductivist views of human culture—views whose genealogy goes back to the reaction against Enlightenment values that increasingly characterized ethnology during the second half of the nineteenth century. This was the same period when the fiction emerged that transmissible "units of culture" were the predominant, if not sole, form of cultural reproduction (e.g., Ratzel 1882–1891). Although Andersson and Read do not address this explicitly, it is clear that their (broadly successful) attempt to downgrade Henrich's "treadmill model" from law to a not-uninteresting speculation with possible utility in modeling some instances of cultural change is an effort also directed against a baleful re-emergence of a tacit essentialism.

Tasmania is key here, as its ostensible archaeology and ethnography were referred to by Henrich in developing his original thesis (2004). He described a Tasmanian aboriginal culture, after the postglacial separation of the territory from mainland Australia, characterized by "severe" and "maladaptive losses of particular kinds of skills and related technologies" (Henrich 2004:197). For example, "despite their cool maritime climate, the Tasmanians... appear to have lost the ability to make cold-weather clothing—a skill that likely allowed them to weather the last glacial maximum" (Henrich 2004:198). But if this ever bothered the Tasmanians, they did not let on. Henry Ling Roth, who, with additional contributors, produced the most comprehensive ethnographic synthesis (1899), notes that the population were a source of wonder to Europeans in terms of the good health

they enjoyed in a cold, wet climate: their dentition was considered near perfect and their skins unblemished (except by outbreaks of disease plausibly attributed to European contact); insulated by red ochre skin applications, a uniquely efficient fire technology, and a well-chosen diet, they avoided damp clothes and rheumatism and slept happily naked under the stars.

Henrich's publication was heavily biased, not only in its failure to cite the principal sources on Tasmanian ethnography (such as Ling Roth et al. 1899), but also in its privileging of the polarized archaeological views of Rhys Jones (e.g., Jones 1977). Jones had already been heavily criticized in his interpretations by Richard Cosgrove and coworkers, in work also ignored by Henrich (e.g., Cosgrove 1999; Holdaway and Cosgrove 1997; see now Cosgrove 2014; and for the broader background of diachronic cultural complexity and adaptation in Australasia, see Hiscock 2008). Jones had followed Tylor in seeing Tasmanian aborigines as "representative of the Paleolithic age" (Cosgrove 1999:359), but the fact that material cultural complexity in these small-scale communities had been further reduced after Tasmania became geographically isolated was seen as a proof of the "treadmill model."

The counter idea, that Tasmanians at contact were legatees of a material culture which they themselves had consciously refined in the most literal sense, choosing a path of expedience over a more costly and risky increase in entailment, is something which (having argued it myself) I am obviously glad to see positively commented on (Taylor 2010). My alternative account was part of a broader argument that rejected the "dual inheritance" theory of hominin evolution out of hand, and I have since sketched more of the background to that rejection (Taylor 2012). Central here is the idea of essentialist "units" of culture—such as "making fire"—that the next generation passively imitates.

Tasmanians, while they could probably make fire *ab initio*, Polynesian style (Backhouse Walker 1900:69–70), typically made it from a curated flame—a fire log or fire stick that the group always carried (with no danger to clothing, as they were naked). When that went out, they either found more (forest fires, for example) or, especially in the wet season, negotiated for it (not giving fire as a gift, even among actively feuding groups, was a universal taboo, and fire was the first gift they gave to the French; see Taylor 2010:143 with references). The "fire-making skill" was thus not a single indivisible transferable entity, although the preferred mode (curation) was the one that produced fire most reliably and quickly. Similar intentional considerations can be argued to underpin all the other so-called deficits in Tasmanian culture, and whether such arguments are accepted as valid or not, they nevertheless demonstrate that our data are underdetermined when not downright imponderable. Such ambiguity is, as Andersson and Read demonstrate, fatal for the strong, law-like version of the "treadmill model."

When V. Gordon Childe wrote that "the environments to which societies are adjusted are worlds of ideas that differ

not only in extent and content, but also in structure" (Childe 1949:22), he based his judgement on knowledge of a very broad range of cases. These included those where large-scale, complex societies had effectively put the brakes on technological development primarily for ideological reasons. The centralized control of bronze production in the Near East, he argued, had stifled progress, while the tribal societies of the central European Bronze Age raced ahead, producing an astounding range of complex metallurgical and metallic innovations. States tended toward theocracy and were prone to stagnation in the realm of material innovation in a way that small-scale and flexible social formations were not. On the other hand, states explicitly prided themselves on producing the material conditions believed to please their gods; that is, their elite members viewed themselves as "skilled" in this connection. In such a hall of mirrors, attempting to account for differences in cultural complexity using a neutral algebra is, at best, limiting.

Krist Vaesen

Philosophy and Ethics, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands; and Human Origins Group, Faculty of Archaeology, University of Leiden, Einsteinweg 2, 2333 CC Leiden, The Netherlands (k.vaesen@tue.nl). 24 VIII 15

Andersson and Read's article is a welcome antidote to the uncritical enthusiasm for the models of Henrich and of Powell and colleagues. Since I am in strong agreement with pretty much everything Andersson and Read write, I will just try to strengthen their case here and offer three additional arguments against the treadmill model.

The first relates to the ethnographic data. Andersson and Read are entirely right that the Oswalt, Inuit, and Oceanic Island data sets do not (Oswalt and Inuit) or only weakly (Oceanic Island) support a relationship between population size and complexity of hunter-gatherer tool kits. But even if a (significant) relationship would be found, that would by no means imply population size to be a driver of cultural complexity. To confirm such a causal relationship, one should (minimally) observe an association between demographic and cultural change—an association which, by definition, cannot be inferred from data sets such as those just mentioned, which contain population and complexity numbers taken at only one point in time—for a correlation between absolute numbers is perfectly consistent with complexity being driven by nondemographic factors (e.g., social or cognitive innovations and adaptivity to environmental conditions), with population size acting merely as a passive constraint.

As to the second addition, Andersson and Read critically review the evidence regarding the Tasmanian case. Although that evidence is of the right kind (i.e., it concerns a possible association between demographic and cultural change), An-

dersson and Read convincingly argue that the archaeological data do not license us to construe the case as one of cultural decline. But it is just as doubtful that demographic change was involved. While Henrich assumes that, before the sea level rise of the Holocene, the foragers inhabiting what is now Tasmania formed a pool of interacting social learners with groups from what is now Bass Strait and mainland Australia, there is no archaeological evidence to suggest any social connections between the regions that would become the mainland and Tasmania. For example, no exotic Tasmanian artifact raw materials, such as Darwin glass, brecciated chert, or blue chert of Late Pleistocene age, have ever been found in Victorian mainland sites dated to the same period (Cosgrove 2015; Cosgrove et al. 2014; Hewitt and Allen 2010; K. Vaesen, M. Collard, W. Roebroeks, and C. Cosgrove, unpublished manuscript, 2015).

A third challenge comes from evidence concerning the Upper Paleolithic transition (as targeted by Powell and colleagues). One may reasonably wonder whether the remarkable cultural developments of the Late Pleistocene really correspond to an increase in cultural complexity—or more specifically, following Powell and colleagues, to an increase in transmission inaccuracy. Yet, even if this were so, a direct association with demography seems ill-supported. Confronting population estimates (adopted by Powell, Shennan, and Thomas [2009, 2010] from Atkinson, Gray, and Drummond [2008]) with estimated dates for the arrival of fully modern human behavior (FMH) in various parts of the world (estimates by Powell, Shennan, and Thomas [2009, 2010]) yields quite a few nontrivial anomalies. Concerning Sub-Saharan Africa, populations grow steadily from ~160 kya onward, yet FMH appears only around 90–75 kya, and FMH disappears, despite population growth, between 75 and 40 kya. Population growth in North and Central Asia starts ~55 kya, whereas the first elements of FMH (namely microliths) emerge only ~43 kya, and FMH in full evolves only ~22 kya. Southern Asian populations increase very markedly 55–45 kya, after which they stabilize; it is in the latter period, not during expansion, that FMH gradually develops. In Australia, FMH arrives fairly suddenly ~20 kya, much after the pronounced population increase 50–45 kya. Another type of anomaly concerns events after the arrival of FMH. In Sub-Saharan Africa, Europe, North and Central Asia, and the Middle East and North Africa, populations continue to expand, whereas they tend to stabilize in Southern Asia and Australia. To salvage their model, Powell, Shennan, and Thomas thus must demonstrate that complexity further increased in the former parts of the world, while Southern Asia and Australia went through a period of cultural stasis.

Powell and colleagues argue that some of these anomalies may be due to the low resolution of single-locus coalescent inferences (i.e., the method used by Atkinson, Gray, and Drummond 2008). However, while a recent multilocus study by Schiffels and Durbin (2014) resolves some of said anomalies, it gives rise to a set of new ones. In Africa, FMH would

first appear at a time when populations were shrinking. In Europe, FMH would arrive at a historic low. Furthermore, population curves for Asia and Europe follow a trajectory that is almost identical, which is at odds with the variation between these two regions with regard to the timing of the appearance of FMH.

To conclude, let me be clear that I do not deny that demography may have a bearing on the mode and tempo of cultural evolution. I merely claim that cultural evolutionists have been too quick in identifying the mechanisms underlying this causal relation, an error that gives rise precisely to the failed predictions described above. On a more positive note, I am convinced, but cannot argue here, that a more promising approach is to revive a tradition that currently seems to have fallen out of favor, namely the tradition set in motion by Thomas Malthus and Ester Boserup.

Acknowledgments

Research by Krist Vaesen was supported by the Dutch Organisation for Scientific Research.

Reply

We are pleased with the positive endorsement of our article by all of the commentators. This gives us the opportunity to expand our article in the directions that they introduce in their comments. We find it striking that all reach essentially the same point, although from a variety of directions: namely, that the models (formal and informal) aimed at accounting for change in cultural complexity need to be restructured to bring them into accord with relevant ethnographic research. In our reply, we first make a few observations about their comments, and then we consider whether, as Gilligan suggests, we should have rejected the treadmill model outright. In so doing, we are led to discuss further the pervasive conceptual problem with the treadmill model introduced by having the same parameter, α , relating sometimes to skillfulness and at other times to complexity.

With regard to the comments, Eriksson, the first commentator, focuses on the fact that good mathematical modeling requires the modeling to be grounded in rigorous ethnographic research that delineates the properties of the processes being modeled mathematically. All too often, he notes, and especially when modeling cultural evolution, the emphasis is on the dynamics encompassed within a model but without having first established the connection between empirical observations and the theoretical processes incorporated in the model. Without a solid empirical foundation, he indicates, models such as the treadmill model are premature. When the empirical connection is found to be wanting, as has been documented with the treadmill model, all too often the response of the modelers and

their supporters is to assume the disconnect is with the empirical observations, rather than using the disconnect as the impetus for rethinking and revising the mathematical model to bring it into accord with empirical observation.

Lack of agreement between empirical observation and model leads Gilligan to suggest that we have not gone far enough in our critique of the treadmill model—that it should be rejected outright given the disconnect between empirical observation and the predictions of the model. Gilligan points out that the prediction of maladaptation as an evolutionary outcome, which is a key conclusion of the treadmill model, is highly problematic, because the loss of technology previously needed for survival, for example, would lead to the demise of a group, not to reduction in complexity. Gilligan points out that the Tasmanians would have died quickly through hypothermia were they to have had the maladaptive loss of bone points and protective clothing predicted by the treadmill model. Here Gilligan makes the connection between empirical observation and what should be modeled that Erikson finds wanting in the treadmill model, as well as pointing out the inadequate consideration of the interplay between cultural evolution and biological evolution in the treadmill model. Hence his conclusion that it should be rejected out of hand. We do not disagree with the thrust of his argument; our interest was to explore whether it might be possible to salvage something from the treadmill model regarding the role of demography in cultural evolution.

In a similar vein, O'Brien and Bentley find a disconnect between empirical data and the theoretical assumption of the treadmill model in its presumption that demography is the prime mover of cultural complexity. While recognizing that demography is, in some sense, associated with cultural complexity—trivially, the material and technological complexity of Western societies depends upon having large populations—they note that demography is not the driver of cultural complexity in the manner derived through the mathematical formalism of the treadmill model. They, too, see the need to connect empirical observation with hypothesized theoretical processes, and they pose the important question, in the context of Paleolithic cave art, of whether it is the process of transmission through imitation underlying the treadmill model that accounts for the coherence and continuity of cultural phenomena like the cave art over many generations, or whether the continuity arises from the systemic and organizational aspect of those phenomena (see chapter 3 in Leaf and Read 2012*b*). In effect, they are questioning whether we better understand cultural evolution through change in the frequency of traits over a population, as is assumed by the dual inheritance model of evolution, or through change at the organizational level of cultural phenomena (see discussions in Andersson, Törnberg, and Törnberg 2014; Lane et al. 2009; Read, Lane, and van der Leeuw 2009).

Taylor also takes up the problem with reducing culture to individual traits. Reducing cultural phenomena in this manner, he notes, has roots going back to the last part of the

nineteenth century. Although this simplifies incorporation of cultural phenomena under the umbrella of Darwinian population genetic models (with transmission expanded to include phenotypic transmission, as in the dual inheritance model of evolution), the cost is denial of the systemic nature of cultural phenomena, as pointed out by Taylor through his reference to Childe's view of societies as being adapted to environments consisting of structured "worlds of ideas" (1949:22; see also Leaf and Read 2012*b*). Taylor brings the "world of ideas" back to the ethnographic level of the Tasmanians through observing that fire, one of the areas in which Tasmanians supposedly were culturally deficient, was integrated into group-level dynamics through being curated and passed on to others, like Mauss's gift, rather than constantly being made anew through the skill (and skillfulness) of individuals, as is assumed in the treadmill model.

Finally, and in keeping with Erikson's argument regarding the need to connect empirical observation with theoretical construct, Vaesen centers his critique on three ways the treadmill model fails to make this link. First is the problematic use of synchronic data to provide empirical support for what is theorized to be a diachronic process. Second, the model requires correlation between decline in population size and the supposed decline in cultural complexity among the Tasmanians, but the archaeological data do not support the hypothesized large population composed of a combined Tasmanian-mainland population interacting before the separation of Tasmania from the mainland by rising seas. The initial large population is required for the model prediction of decrease in cultural complexity with reduction of population size. Third, the worldwide pattern during the Pleistocene regarding change in population size and cultural complexity, rather than showing that increased cultural complexity (measured by the appearance of fully modern human behavior) and population size varies in parallel, as the treadmill model predicts, shows the contrary, with both episodes of population decline and increasing complexity and episodes of population increase but no increase in complexity. Vaesen ends his comment where Erikson began, by regarding the need to connect ethnographic observation with theory, suggesting that it may be more useful to develop arguments about cultural change, beginning with the interplay between demography and agricultural intensification discussed by the economist Ester Boserup (1965).

With regard to Gilligan's observation that the treadmill model should be rejected out of hand, we want to remark that we actually do this for its most central intended explanatory service: as a formal model for change in cultural complexity. That the model is not universally right, though, does not mean that it is universally wrong. Details of model formulation aside, the intuition that evolution includes an opposition between creative and destructive forces appears too robust, for example, to dismiss easily. We suggest that this opposition may be more evident in small ($n < 15$), rather than large, groups. Andersson and Törnberg (2016) have found support for this idea by modeling small group, causal microdynamics.

However, models—even when cast in the language of cultural traits, cultural learning, and interacting social learners—do not clarify the relationship between cultural complexity and number of interacting social learners when the model (see Henrich 2009) is equivalent to showing the well-known result from population genetics that the equilibrium trait frequency is driven, except for low selection values, primarily by selection intensity (the number of social learners), or the model (see Henrich 2016:213–214) is equivalent to showing that mutation with (positive) selection is a vastly faster evolutionary process for increasing trait frequency than neutral mutations (inventions without social learning). Demography does play a role in the dynamics of cultural complexity; the challenge is to work out rigorously what that role is.

As we point out, the formal treadmill model models only change in skillfulness, not complexity, and assumes complexity is constant. Hence, if we want a formal model or models for the processes by which complexity changes, we must discard the formal treadmill model entirely and begin anew (possibly using its basic logic as one of several components). Furthermore, the confusion introduced by having α interpreted first as loss of skillfulness under imitation in the formal model, but then as a measure of complexity in the informal model, needs to be removed. This confusion between whether α measures skillfulness or complexity was noted in Read's (2006) comment that Henrich (2004) confounds the knowledge needed to do a task (i.e., complexity) with the level of motor skills that individuals develop for doing a task (i.e., skillfulness). Furthermore, Henrich does not take into account the biological dimension (as noted by Gilligan) involved in skillfulness; hence, the "imitation model of skill transmittal, rather than being a general model as proposed by Henrich, may simply be a special case" (Read 2006:165).

Missing in the treadmill model is differentiation (in the case of artifacts) between transmission of performance (relating to skillfulness) and concept and design (relating to complexity). Performance is transmitted through learning and experience (including imitation) and, in a developmental sense, involves going from having limited skillfulness to having the skillfulness needed for the task at hand. This is where variation at the individual level and how it relates to biological evolution come into play. Furthermore, skillfulness does not increase without bound, and it may be voluntarily limited according to the level of skillfulness needed for doing the task. In contrast, transmission of the concept and design modality involves transmission through communication and incorporates invention and innovation, with the former occurring at the individual level and the latter relating to what is known and held in common by group members and how that may change and diffuse through a group (see Rogers 2003). In other words, the concept and design modality is part of the cultural knowledge transmitted through enculturation, with the latter a life-long process that, in its initial stages, involves interaction of children with members of their community and with their parents. Invention and innovation are not determined by im-

itation (although invention may occur when an individual is learning skills through imitation), and adaptation comes into play as a process through which commonly recognized problems (such as ensuring successful procurement of food resources) are addressed individually and collectively, such that inventions become innovations (Van Pool, O'Brien, and Lyman 2015).

Our sense is that the claim that population size is a primary driver for the emergence of complex cultural systems has been too easily accepted, because it resounds strongly with intuitions based on the evolutionary trajectory of more recent (Holocene) societies and how this relates to change in cultural complexity. Yet even here, population size becomes important in entirely different ways than modeled in the treadmill model. As discussed by Boserup (1965), elaboration in the division of labor for the production of material goods and food resources, coupled with increasing food production through labor intensification, also gave impetus to inventions and innovations whose implementation (performance) required, in turn, larger labor units and networks of interacting labor units as well as an increasingly hierarchical organizational structure. Taking into account change at the level of the organizational structure of societies takes us far away from the treadmill story, even within small-scale societies. Consider, for example, the sealing partner system of the Netsilik Inuit that was necessary for procuring seals through the Arctic ice in the winter (Balicki 1970; Read 2005); until that system of sealing partners could be put together, the Arctic coast could not be occupied under the climate conditions faced by the Netsilik. Not population size but innovation at the level of the organization of a social system was the driving factor for developing the complex system of sealing partners.

—Dwight Read and Claes Andersson

References Cited

- Allen, H. 1979. Left out in the cold: why the Tasmanians stopped eating fish. *Artefact* 4:1–10.
- Andersson, C. 2011. Palaeolithic punctuations and equilibria: did retention rather than invention limit technological evolution? *PalaeoAnthropology* 2011:243–259.
- . 2013. Fidelity and the emergence of stable and cumulative socio-technical systems. *PalaeoAnthropology* 2013:88–103.
- Andersson, C., and D. W. Read. 2014. Group size and cultural complexity. *Nature* 511:E1.
- Andersson, C., A. Törnberg, and P. Törnberg. 2014. An evolutionary developmental approach to cultural evolution. *Current Anthropology* 55(2): 154–174.
- Andersson, C., and P. Törnberg. 2016. Fidelity and the speed of the treadmill: the combined impact of population size, transmission fidelity and selection on cultural complexity. *American Antiquity* 81(3).
- Arthur, W. 2011. *Evolution: a developmental approach*. Chichester: Wiley-Blackwell.
- Atkinson, Q., R. Gray, and A. Drummond. 2008. mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. *Molecular Biology and Evolution* 25:468–474. [KV]
- Atkinson, C., M. J. O'Brien, and A. Mesoudi. 2012. Adult learners in a novel environment use prestige-biased social learning. *Evolutionary Psychology* 10:519–537. [MJO/RAB]

- Backhouse Walker, J. 1900. *The Tasmanian aborigines*. Hobart, Tasmania: Vail. [TT]
- Baldini, R. 2013. Revisiting the effect of population size on cumulative cultural evolution. *bioRxiv*. doi: <http://dx.doi.org/10.1101/001529>.
- Balikci, A. 1970. *The Netsilik Eskimo*. Doubleday.
- Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson. 2013. The diffusion of microfinance. *Science* 341:1236498. [MJO/RAB]
- Bell, A. V. 2014. Cultural evolution and the way we count. *Proceedings of the National Academy of Sciences of the USA* 111(4):1227–1228.
- . 2015. Linking observed learning patterns to the evolution of cultural complexity. *Current Anthropology* 56(2):277–281.
- Bentley, R. A., and M. J. O'Brien. 2011. The selectivity of social learning and the tempo of cultural evolution. *Journal of Evolutionary Psychology* 9(2): 125–141.
- . 2015. Collective behavior, uncertainty and environmental change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 373:20140461. [MJO/RAB]
- Bentley, R. A., M. J. O'Brien, and W. A. Brock. 2014. Mapping collective behavior in the big-data era. *Behavioral and Brain Sciences* 37:63–119.
- Bettencourt, L., and G. West. 2010. A unified theory of urban living. *Nature* 467:912–913.
- Boserup, E. 1965. *The conditions of agricultural growth: the economics of agrarian change under population pressure*. London: Allen & Unwin.
- Boswell, J. 1848 [1791]. *The life of Samuel Johnson, L.L.D., including a journal of his tour to the Hebrides*. Vol. 5. London: Murray. [MJO/RAB]
- Bowdler, S., and H. Lourandos. 1982. Both sides of Bass Strait. Pp. 121–132 in S. Bowdler, ed. *Coastal archaeology in eastern Australia*. Canberra: Australian National University.
- Bowern, C. 2012. The riddle of Tasmanian languages. *Proceedings of the Royal Society B: Biological Sciences* 279:4590–4595. [IG]
- Boyd, R., and P. J. Richerson. 1985. *Culture and the evolutionary process*. Chicago: University of Chicago Press.
- Bril, B., V. Roux, and G. Dietrich. 2005. Stone knapping: Khambhat (India), a unique opportunity? Pp. 53–72 in V. Roux and B. Bril, eds. *Stone knapping: the necessary conditions for a uniquely hominin behaviour*. Cambridge: McDonald Institute for Archaeological Research.
- Buc, N. 2010. Experimental series and use-wear in bone tools. *Journal of Archaeological Science* 38:546–557.
- Caldwell, C. A., and A. E. Millen. 2010. Human cumulative culture in the laboratory: effects of (micro) population size. *Learning and Behavior* 38 (3):310–318.
- Cavalli-Sforza, L. L., and M. W. Feldman. 1981. *Cultural transmission and evolution*. Stanford, CA: Stanford University Press.
- Centola, D., and A. Baronchelli. 2015. The spontaneous emergence of conventions: an experimental study of cultural evolution. *Proceedings of the National Academy of Sciences* 112:1989–1994. [MJO/RAB]
- Chaisson, E. 2014. Practical applications of cosmology to human society. *Natural Science* 6:767–796.
- Childe, V. G. 1949. *Social worlds of knowledge*. London: Oxford University Press.
- Cipriani, L. 1966. *The Andaman Islanders*. New York: Praeger. [IG]
- Claidière, N., and D. Sperber. 2010. Imitation explains the propagation, not the stability of animal culture. *Proceedings of the Royal Society B: Biological Sciences* 277(1681):651–659.
- Cohen, M. 1985. Prehistoric hunter-gatherers: the meaning of social complexity. Pp. 99–119 in T. D. Price and J. A. Brown, eds. *Prehistoric hunter-gatherers: the emergence*. Waltham, MA: Academic Press.
- Colhoun, E. A. 2000. Vegetation and climate change during the Last Interglacial-Glacial cycle in western Tasmania, Australia. *Palaeogeography, Palaeoclimatology, Palaeoecology* 155:195–209. [IG]
- Colebrooke, R. H. 1807. On the Andaman Islands. *Asiatic Researches* 4:385–394. [IG]
- Collard, M., B. Buchanan, J. Morin, and A. Costopoulos. 2011. What drives the evolution of hunter-gatherer subsistence technology? a reanalysis of the risk hypothesis with data from early contact era populations in the Pacific Northwest. *Philosophical Transactions of the Royal Society B* 366: 1129–1138.
- Collard, M., B. Buchanan, and M. J. O'Brien. 2013. Population size as an explanation for patterns in the Paleolithic archaeological record. *Current Anthropology* 54:S388–S396.
- Collard, M., B. Buchanan, M. J. O'Brien, and J. Scholnick. 2013a. Risk, mobility or population size? drivers of technological richness among contact-period western North American hunter-gatherers. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368:20120412.
- Collard, M., M. Kemery, and S. Banks. 2005. Causes of toolkit variation among hunter-gatherers: a test of four competing hypotheses. *Canadian Journal of Archaeology* 29:1–19.
- Collard, M., A. Ruttle, B. Buchanan, M. J. O'Brien. 2012. Risk of resource failure and toolkit variation in small-scale farmers and herders. *PLoS ONE* 7(7):e40975. doi:10.1371/journal.pone.0040975.
- Collard M., A. Ruttle, B. Buchanan, and M. J. O'Brien. 2013b. Population size and cultural evolution in nonindustrial food-producing societies. *PLoS ONE* 8(9), e72628. doi:10.1371/journal.pone.0072628.
- Cooper, Z. 1989. Analysis of the nature of contacts with the Andaman Islands during the last two millennia. *South Asian Studies* 5:133–147. [IG]
- Cosgrove, C., A. Pike-Tay, and W. Roebroeks. 2014. Tasmanian archaeology and reflections on modern human behaviour. Pp. 175–188 in R. Dennell and M. Porr, eds. *Southern Asia, Australia and the search for human origins*. Cambridge: Cambridge University Press. [KV, TT]
- Cosgrove, R. 1999. Forty-two degrees south: the archaeology of Late Pleistocene Tasmania. *Journal of World Prehistory* 13(4):357–402. [TT]
- . 2015. Raw material movement and past Tasmanian Aboriginal interaction: implications for understanding isolated human population dynamics. *Archaeology in Oceania* 50:69–81. [KV]
- Cosgrove, R., and J. Allen. 2001. Prey choice and hunting strategies in the Late Pleistocene: evidence from southwest Tasmania. Pp. 397–429 in A. Anderson, I. Lilley, and S. O'Connor, eds. *Histories of old ages: essays in honour of Rhys Jones*. Canberra, Australia: Pandanus, Research School of Pacific and Asian Studies, Australian National University.
- Csibra, G., and G. Gergely. 2009. Natural pedagogy. *Trends in Cognitive Sciences* 13(4):148–153.
- . 2011. Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366:1149–1157.
- Darwin, C. R. 1839. Journal and remarks, 1832–1836. In R. Fitz-Roy. *Narrative of the surveying voyages of His Majesty's Ships Adventure and Beagle, between the years 1826 and 1836, describing their examination of the southern shores of South America, and the Beagle's circumnavigation of the globe*. Vol. III. London: Henry Colburn. [IG]
- Derrida, J. 2002. The animal that therefore I am (more to follow). David Wills, trans. *Critical Inquiry* 28:369–418. [IG]
- Descola, P. 2005. *Beyond nature and culture*. Janet Lloyd, trans. Chicago: University of Chicago Press. [IG]
- Domingo, E., J. Sheldon, and C. Perales. 2012. Viral quasispecies evolution. *Microbiology and Molecular Biology Reviews* 76(2):159–216.
- Dean, L. G., R. L. Kendal, S. J. Schapiro, B. Thierry, and K. N. Laland. 2012. Identification of the social and cognitive processes underlying human cumulative culture. *Science* 335(6072):1114–1118. doi:10.1126/science.1213969.
- Dean, L. G., G. L. Vale, K. N. Laland, E. Flynn, and R. L. Kendal. 2014. Human cumulative culture: a comparative perspective. *Biological Reviews of the Cambridge Philosophical Society* 89(2):284–301. doi:10.1111/brv.12053.
- Derex, M., M.-P. Beugin, B. Godelle, and M. Raymond. 2013. Experimental evidence for the influence of group size on cultural complexity. *Nature* 503 (7476):389–391.
- . 2014. Derex et al. reply. *Nature* 511(7507):E2.
- Drake, J. W., B. Charlesworth, D. Charlesworth, and J. F. Crow. 1998. Rates of spontaneous mutation. *Genetics* 148(4):1667–1686.
- Eerkens, J. W., and C. P. Lipo. 2005. Cultural transmission, copying errors, and the generation of variation in material culture and the archaeological record. *Journal of Anthropological Archaeology* 24:316–334. [MJO/RAB]
- Eigen, M. 1971. Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58(10):465–523.
- Eigen, M., and P. Schuster. 1977. The hypercycle: a principle of natural self-organization. *Naturwissenschaften* 64:541–565.
- Engerman, S. L. 1997. The standard of living debate in international perspective: measures and indicators. Pp. 17–45 in R. H. Steckel and R. Floud, eds. *Health and welfare during industrialization*. Chicago: University of Chicago Press. [MJO/RAB]
- Enquist, M., P. Strimling, K. Eriksson, K. N. Laland, J. Sjöstrand. 2010. One cultural parent makes no culture. *Animal Behaviour* 79(6):1353–1362.
- Erwin, D. H., and J. W. Valentine. 2013. *The Cambrian explosion*. Greenwood Village, CO: Roberts.
- Flannery, K. V. 1969. Origins and ecological effects of early domestication in Iran and the Near East. Pp. 73–100 in P. J. Ucko and G. W. Dimbleby, eds. *The domestication and exploitation of plants and animals*. Chicago: Aldine.

- . 1976. *The early Mesoamerican village*. New York: Academic Press.
- Gamble, C. 2007. *Origins and revolutions: human identity in earliest prehistory*. Cambridge, MA: Cambridge University Press.
- . 2010. Technologies of separation and the evolution of social extension. Pp. 17–42 in R. I. M. Dunbar, C. Gamble, and J. Gowlett, eds. *Social brain, distributed mind*. Oxford: Oxford University Press.
- Geels, F., and J. Schot. 2007. Typology of sociotechnical transition pathways. *Research Policy* 36(3):399–417.
- Geels, F. W. 2002. Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. *Research Policy* 31:1257–1274.
- Gilligan, I. 2007a. Clothing and modern human behaviour: prehistoric Tasmania as a case study. *Archaeology in Oceania* 42(3):102–111.
- . 2007b. *Another Tasmanian paradox: clothing and thermal adaptations in Aboriginal Australia*. BAR International Series 1710. Oxford: Archaeopress. [IG]
- . 2010. The prehistoric development of clothing: archaeological implications of a thermal model. *Journal of Archaeological Method and Theory* 17:15–80. [IG]
- . 2014. Clothing and modern human behaviour: the challenge from Tasmania. Pp. 287–301 in R. Dennell and M. Porr, eds. *Southern Asia, Australia and the search for human origins*. New York: Cambridge University Press. [IG]
- Gilligan, I., and D. Bulbeck. 2007. Environment and morphology in Australian Aborigines: a re-analysis of the Birdsell database. *American Journal of Physical Anthropology* 134:75–91. [IG]
- Gilligan, I., S. Chandraphak, and P. Mahakkanukrauh. 2013. Femoral neck-shaft angle in humans: variation relating to climate, clothing, lifestyle, sex, age and side. *Journal of Anatomy* 223:133–151. [IG]
- Gould, S. J., and R. C. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceeding of the Royal Society of London B: Biological Sciences* 205:581–598.
- Hamilton, M. J., and B. Buchanan. 2009. The accumulation of stochastic copying errors causes drift in culturally transmitted technologies: quantifying Clovis evolutionary dynamics. *Journal of Anthropological Archaeology* 28:55–69. [MJO/RAB]
- Henrich, J. 2004. Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses: the Tasmanian case. *American Antiquity* 69(2):197–214.
- . 2006. Understanding cultural evolutionary models: a reply to Read's critique. *American Antiquity* 71(4):771–782.
- . 2009. The evolution of innovation-enhancing institutions. Pp. 99–120 in S. J. Shennan and M. J. O'Brien, eds. *Innovation in cultural systems: contributions in evolution anthropology*. Cambridge, MA: MIT Press.
- . 2016. *The secret of our success: how culture is driving human evolution, domesticating our species and making us smarter*. Princeton, NJ: Princeton University Press.
- Henrich, J., and R. Boyd. 2002. On modeling cognition and culture: why cultural evolution does not require replication of representations. *Journal of Cognition and Culture* 2(2):87–111.
- Henrich, J., and J. Broesch. 2011. On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366(1567):1139–1148. doi: 10.1098/rstb.2010.0323.
- Hewitt, G., and J. Allen. 2010. Site disturbance and archaeological integrity: the case of Bend Road, an open site in Melbourne spanning pre-LGM Pleistocene to Late Holocene periods. *Australian Archaeology* 70:1–16. [KV]
- Hidalgo, C. A., and R. Hausmann. 2009. The building blocks of economic complexity. *Proceedings of the National Academy of Sciences of the USA* 106:10570–10575. [MJO/RAB]
- Hill, K. R., B. M. Wood, J. Baggio, A. M. Hurtadom, and R. T. Boyd. 2014. Hunter-gatherer inter-band interaction rates: implications for cumulative culture. *PLoS One* 9(7):e102806. doi: 10.1371/journal.pone.0102806.
- Hiscock, P. 2008. *The archaeology of ancient Australia*. London: Routledge.
- Hodder, I. 2012. *Entangled: an archaeology of the relationships between humans and things*. New York: Wiley-Blackwell.
- Holdaway, S., and R. Cosgrove. 1997. The archaeological attributes of behaviour: difference or variability. *Endeavour* 21(2):66–71. [TT]
- Hope, G. S. 1978. The Late Pleistocene and Holocene vegetational history of Hunter Island, north-western Tasmania. *Australian Journal of Botany* 26:493–514. [IG]
- Horner, V., A. Whiten, E. Flynn, and F. B. M. de Waal. 2006. Faithful replication of foraging techniques along cultural transmission chains by chimpanzees and children. *Proceedings of the National Academy of Sciences of the USA* 103(37):13878–13883.
- Hume, D. 1985 [1777]. *Essays, moral, political and literary*. Indianapolis: Liberty Classics. [MJO/RAB]
- Husserl, E. G. 1954. *The crisis of European sciences and transcendental phenomenology: an introduction to phenomenological philosophy*. Translated by David Carr. Chicago: Northwestern University Press, 1970. [IG]
- Ingold, T. 1980. *Hunters, pastoralists and ranchers: reindeer economies and their transformations*. Cambridge: Cambridge University Press. [IG]
- Johnson, A. W., and T. Earl. 2000. *From foraging group to agrarian state*. 2nd ed. Stanford, CA: Stanford University Press.
- Jones, R. 1971. *Rocky Cape and the problem of the Tasmanians*. Sydney: University of Sydney.
- . 1977. The Tasmanian paradox. Pp. 189–204 in R. V. S. Wright, ed. *Stone tools as cultural markers: change, evolution and complexity*. Canberra: Australian Institute of Aboriginal Studies. [IG]
- . 1990. From Kakadu to Kutikina: the southern continent at 18,000 years ago. Pp. 264–295 in O. Soffer and C. Gamble, eds. *Low latitudes, vol. 2 of The world at 18000 BP*. London: Unwin Hyman.
- Keely, L. H. 1988. Hunter-gatherer economic complexity and “population pressure”: a cross-cultural analysis. *Journal of Anthropological Archaeology* 7:373–411.
- Kempe, M., and A. Mesoudi. 2014. An experimental demonstration of the effect of group size on cultural accumulation. *Evolution and Human Behavior* 35(4):285–290.
- Klein, R. 2002. *The dawn of human culture*. New York: Wiley. [MJO/RAB]
- Kline, M. A., and R. Boyd. 2010. Population size predicts technological complexity in Oceania. *Proceedings of the Royal Society B: Biological Sciences* 277(1693):2559–2564.
- Kobayashi, Y., and K. Aoki. 2012. Innovativeness, population size and cumulative cultural evolution. *Theoretical Population Biology* 82(1):38–47.
- Laland, K. N., N. Atton, and M. M. Webster. 2011. From fish to fashion: experimental and theoretical insights into the evolution of culture. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366(1567):958–968.
- Lane, D., and R. R. Maxfield. 2005. Ontological uncertainty and innovation. *Journal of Evolutionary Economics* 15:3–50.
- Lane, D. A. 2011. Complexity and innovation dynamics. Pp. 63–80 in C. Antonelli, ed. *Handbook on the economic complexity of technological change*. Cheltenham, UK: Elgar.
- Lane, D. A., R. R. Maxfield, D. W. Read, and S. E. van der Leeuw. 2009. From population to organization thinking. Pp. 11–42 in D. A. Lane, S. E. van der Leeuw, D. Pumain, and G. B. West, eds. *Complexity perspectives in innovation and social change*. Berlin: Springer.
- Laubichler, M. D., and J. Maienschein. 2013. Developmental evolution. Pp. 375–382 in M. Ruse, ed. *The Cambridge encyclopedia of Darwin and evolutionary thought*. Cambridge: Cambridge University Press.
- Leaf, M., D. Read. 2012a. *The conceptual foundation of human society and thought: anthropology on a new plane*. Lanham, MD: Lexington.
- . 2012b. *Human thought and social organization: anthropology on a new plane*. Lanham, MD: Lexington.
- Lewis, H. M., and K. N. Laland. 2012. Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1599):2171–2180.
- Ling Roth, H., M. E. Butler, J. Bvackhouse Walker, and J. G. Garson. 1899. *The aborigines of Tasmania*. 2nd ed. Halifax, UK: King. [TT]
- Malakoff, D. 2013. Are more people necessarily a problem? *Science* 333:544–546. [MJO/RAB]
- Mesoudi, A. 2011a. Variable cultural acquisition costs constrain cumulative cultural evolution. *PLoS ONE* 6(3):15–17.
- . 2011b. *Cultural evolution*. Chicago: University of Chicago Press.
- Mithen, S. J. 1996. *The prehistory of the mind: the cognitive origins of art, religion, and science*. London: Thames & Hudson. [MJO/RAB]
- Mouat, F. J. 1863. *Adventures and researches among the Andaman Islanders*. London: Hurst & Blackett. [IG]
- Mulvaney J., and J. Kamminga. 1999. *Prehistory of Australia*. Washington, DC: Smithsonian Institution.
- Muthukrishna, M., B. W. Shulman, V. Vasilescu, and J. Henrich. 2014. Sociality influences cultural complexity. *Proceeding of the Royal Society of London B: Biological Sciences* 281(1774):20132511.
- Nakahashi, W. 2014. The effect of cultural interaction on cumulative cultural evolution. *Journal of Theoretical Biology* 352:6–15. doi:10.1016/j.jtbi.2014.02.032.

- Newcomer, M. 1974. Study and replication of bone tools from Ksar Akil (Lebanon). *World Archaeology* 6:138–153.
- Noetling, F. 1911. The food of the Tasmanian Aborigines. *Papers and Proceedings of the Royal Society of Tasmania* 63:279–305.
- O'Brien, M. J., and R. A. Bentley. 2011. Stimulated variation and cascades: two processes in the evolution of complex technological systems. *Journal of Archaeological Method and Theory* 18(4):309–335. doi:10.1007/s10816-011-9110-7.
- Odling-Smee, F. J., K. N. Laland, and M. W. Feldman. 2003. *Niche construction: the neglected process in evolution*. Monographs in Population Biology 37. Princeton, NJ: Princeton University Press.
- Oswalt, W. H. 1976. *An anthropological analysis of food-getting technologies*. New York: Wiley.
- Pan, W., G. Ghoshal, C. Krumme, M. Cebrian, and A. Pentland. 2013. Urban characteristics attributable to density-driven tie formation. *Nature Communications* 4:1961. [MJO/RAB]
- Popper, K. R. 1972. *Objective knowledge: an evolutionary approach*. Oxford: Oxford University Press. [IG]
- Powell, A., S. Shennan, and M. G. Thomas. 2009. Late Pleistocene demography and the appearance of modern human behavior. *Science* 324:1298–1301.
- . 2010. Demography and variation in the accumulation of culturally inherited skills. Pp. 137–160 in M. J. O'Brien and S. Shennan, eds. *Innovation in cultural systems: contributions from evolutionary anthropology*. Cambridge, MA: MIT Press.
- Price, G. R. 1970. Selection and covariance. *Nature* 227:520–521.
- Pradhan, G. R., C. Tennie, and C. P. van Schaik. 2012. Social organization and the evolution of cumulative technology in apes and hominins. *Journal of Human Evolution* 63(1):180–190.
- Querbes, A., K. Vaesen, and W. Houkes. 2014. Complexity and demographic explanations of cumulative culture. *PLoS ONE* 9(7):e102543.
- Ratzel, F. 1882–1891. *Anthropogeographie: Die geographische Verbreitung des Menschen*. Stuttgart: Engelhorn. [TT]
- Read, D. 1987. Foraging society organization: a simple model of a complex transition. *European Journal of Operational Research* 30:320–326.
- . 2001. What is kinship? Pp. 78–117 in R. Feinberg and M. Ottenheimer, eds. *The cultural analysis of kinship: the legacy of David Schneider and its implications for anthropological relativism*. Urbana: University of Illinois Press.
- . 2005. Some observations on resilience and robustness in human systems. *Cybernetics and Systems* 36(8):773–802.
- . 2006. Tasmanian knowledge and skill: maladaptive imitation or adequate technology? *American Antiquity* 71:164–184.
- . 2007. Kinship theory: a paradigm shift. *Ethnology* 46(4):329–364.
- . 2008. An interaction model for resource implement complexity based on risk and number of annual moves. *American Antiquity* 73(4):599–625.
- . 2009. The misuse of a mathematical model: the Tasmanian case (reply to Henrich's response). *UCLA: Human Complex Systems*. <http://repositories.cdlib.org/hcs/DWR2009A>.
- . 2010. From experiential-based to relational-based forms of social organization: a major transition in the evolution of *Homo sapiens*. Pp. 203–234 in R. Dunbar, C. Gamble, and J. Gowlett, eds. *Social brain, distributed mind*. Oxford: Oxford University Press.
- . 2012a. *How culture makes us human: primate social evolution and the formation of human societies*. Walnut Creek, CA: Left Coast.
- . 2012b. Population size does not predict artifact complexity: analysis of data from Tasmania, Arctic hunter-gatherers, and Oceania. *UCLA: Human Complex Systems*. <http://www.escholarship.org/uc/item/61n4303q>.
- . 2013. Reconstructing the proto-Polynesian terminology: kinship terminologies as evolving logical structures. Pp. 59–91 in P. McConnell, I. Keen, and R. Hendery, eds. *Kinship systems: change and reconstruction*. Salt Lake City: University of Utah Press.
- Read, D., D. A. Lane, and S. E. van der Leeuw. 2009. The innovation innovation. Pp. 43–84 in D. A. Lane, S. E. van der Leeuw, D. Pumain, and G. B. West, eds. *Complexity perspectives in innovation and social change*. Berlin: Springer.
- Read, D., and S. LeBlanc. 2003. Population growth, carrying capacity, and conflict. *Current Anthropology* 44(1):59–85.
- Richerson, P. J. 2013. Group size determines cultural complexity. *Nature* 503: 351–352.
- Richerson, P. J., R. Baldini, A. Bell, K. Demps, K. Frost, V. Hillis, S. Mathew, et al. 2016. Cultural group selection plays an essential role in explaining human cooperation: a sketch of the evidence. *Behavioral and Brain Sciences* 39:e30.
- Rogers, E. 2003. *Diffusion of innovations*. 5th edition. New York: Free Press.
- Rouse, I. 1939. *Prehistory in Haiti: a study in method*. Yale University Publications in Anthropology 21. New Haven, CT: Yale University Press.
- Sahlins, M. D. 2008. *The Western illusion of human nature, with reflections on the long history of hierarchy, equality, and the sublimation of anarchy in the West, and comparative notes on other conceptions of the human condition*. Chicago: Prickly Paradigm. [IG]
- Schiffels, S., and R. Durbin. 2014. Inferring human population size and separation history from multiple genome sequences. *Nature Genetics* 46:919–925. [KV]
- Schiffer, M. B. 2005. The devil is in the details: the cascade model of invention processes. *American Archaeology* 70(3):485–502.
- Shennan, S. 2001. Demography and cultural innovation: a model and its implications for the emergence of modern human culture. *Cambridge Archeological Journal* 11(1):5–16.
- Shott, M. 1986. Technological organization and settlement mobility: an ethnographic examination. *Journal of Anthropological Research* 42:15–51.
- Speth, J., K. A. Spielmann. 1983. Energy source, protein metabolism, and hunter-gatherer subsistence strategies. *Journal of Anthropological Archaeology* 2:1–31.
- Stephens, D. W., and J. R. Krebs. 1986. *Foraging theory*. Princeton, NJ: Princeton University Press.
- Sterelny, K. 2011. From hominins to humans: how sapiens became behaviourally modern. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366(1566):809–822.
- Stiner, M. C., and S. L. Kuhn. 2006. Changes in the “connectedness” and resilience of Paleolithic societies in Mediterranean ecosystems. *Human Ecology* 34(5):693–712.
- Stiner, M. C., N. D. Munro, and T. A. Surovell. 2000. The tortoise and the hare. *Current Anthropology* 41(1):39–73.
- Taylor, T. 2010. *The artificial ape: how technology changed the course of human evolution*. New York: Palgrave Macmillan.
- Taylor, T. 2012. The problem of “Darwinizing” culture (or memes as the new phlogiston). Pp. 65–82 in M. Brinkworth and F. Weinert, eds. *Evolution 2.0: implications of Darwinism in philosophy and the social and natural sciences*. New York: Springer. [TT]
- Temple, R. C. 1901. An unpublished XVIII Century document about the Andamans. *Indian Antiquary* 30:232–238. [IG]
- Tennie, C., J. Call, and M. Tomasello. 2009. Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1528):2405–2415.
- Texler, J. 1983. *Composition of foods: finfish and shellfish products—raw, processed and prepared*. Agriculture Handbook (United States Department of Agriculture) no. 8–15. Washington, DC: Government Printing Offices.
- Tomasello, M. 1991. Chimpanzee culture? *National Student Speech Language Hearing Association Journal* 18:73–75.
- . 1999. *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Torrence, R. 1983. Time budgeting and hunter-gatherer technology. Pp. 11–22 in G. Bailey, ed. *Hunter-gatherer economy in prehistory: a European perspective*. Cambridge: Cambridge University Press.
- . 1989. Re-tooling: towards a behavioral theory of stone tools. Pp. 57–66 in R. Torrence, ed. *Time, energy and stone tools*. Cambridge: Cambridge University Press.
- . 2000. Hunter-gatherer technology: macro- and microscale approaches. Pp. 99–143 in C. Panter-Brick, R. H. Layton, and P. Rowley-Conwy, eds. *Hunter-gatherers: an interdisciplinary perspective*. Cambridge: Cambridge University Press.
- Tostevin, G. B. 2012. *Seeing lithics*. Oxford: Oxbow.
- Tyzzar E. E. 1936. The “simple bone point” of the shell-heaps of the north-eastern Algonkian area and its probable significance. *American Antiquity* 1:261–279.
- Vaesen, K. 2012a. Cumulative cultural evolution and demography. *PLoS One* 7(7):e40989.
- . 2012b. The cognitive bases of human tool use. *Behavioral and Brain Sciences* 35:203–218. doi:10.1017/S0140525X11001452.
- VanPool, T. L., M. J. O'Brien, and R. L. Lyman. 2015. Innovation and natural selection in Paleoindian projectile points from the American southwest. Pp. 61–82 in N. Goodale and W. Andrefsky, Jr., eds. *Lithic technological systems and evolutionary theory*. New York: Cambridge University Press.
- Wang, H.-W., B. Mitra, T. K. Chaudhuri, M. G. Palanichamy, Q.-P. Kong, and Y.-P. Zhang. 2011. Mitochondrial DNA evidence supports northeast Indian origin of the aboriginal Andamanese in the Late Paleolithic. *Journal of Genetics and Genomics* 38:117–122. [IG]

- Wimsatt, W. C. 2013. Articulating Babel: an approach to cultural evolution. *Studies in History and Philosophy of Biological and Biomedical Sciences* 44:563–571.
- Whiten, A., N. McGuigan, S. Marshall-Pescini, and L. M. Hopper. 2009a. Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1528):2417–2428.
- Whiten, A., K. Schick, and N. Toth. 2009b. The evolution and cultural transmission of percussive technology: integrating evidence from palaeoanthropology and primatology. *Journal of Human Evolution* 57(4):420–435.