

# The Driving Forces of Cultural Complexity

## Neanderthals, Modern Humans, and the Question of Population Size

Laurel Fogarty<sup>1,2</sup> · Joe Yuichiro Wakano<sup>3</sup> ·  
Marcus W. Feldman<sup>1</sup> · Kenichi Aoki<sup>4</sup>

Published online: 25 October 2016

© Springer Science+Business Media New York 2016

**Abstract** The forces driving cultural accumulation in human populations, both modern and ancient, are hotly debated. Did genetic, demographic, or cognitive features of behaviorally modern humans (as opposed to, say, early modern humans or Neanderthals) allow culture to accumulate to its current, unprecedented levels of complexity? Theoretical explanations for patterns of accumulation often invoke demographic factors such as population size or density, whereas statistical analyses of variation in cultural complexity often point to the importance of environmental factors such as food stability, in determining cultural complexity. Here we use both an analytical model and an agent-based simulation model to show that a full understanding of the emergence of behavioral modernity, and the cultural evolution that has followed, depends on understanding and untangling the complex relationships among culture, genetically determined cognitive ability, and demographic history. For example, we show that a small but growing population could have a different number of cultural traits from a shrinking population with the same absolute number of individuals in some circumstances.

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s12110-016-9275-6) contains supplementary material, which is available to authorized users.

✉ Laurel Fogarty  
uccolfo@ucl.ac.uk

<sup>1</sup> Department of Biology, Stanford University, Herrin Labs, 371 Serra Mall, Stanford, CA 94305, USA

<sup>2</sup> CoMPLEX, University College London, Gower Street, London WC1E 6BT, UK

<sup>3</sup> Meiji Institute for Advanced Study of Mathematical Sciences, Meiji University, Nakano 4-21-1, Nakano-ku, Tokyo 164-8525, Japan

<sup>4</sup> Organization for the Strategic Coordination of Research and Intellectual Properties, Meiji University, Nakano 4-21-1, Nakano-ku, Tokyo 164-8525, Japan

**Keywords** Cultural complexity · Population size · Demography · Cultural evolution · Neanderthals · Modern humans

Since migration out of Africa, modern humans have spread to every continent and succeeded under a great variety of ecological conditions. This ecological and demographic success is often attributed to the human capacity to accumulate large amounts of useful information that can be passed between individuals and across generations. However, exactly what features of behaviorally modern humans (and not other hominids) allowed culture to accumulate to this unprecedented level is hotly debated within the fields of archaeology and anthropology (e.g., Klein 2000; Powell et al. 2009; Shennan 2001; Stiner and Kuhn 2006). Identifying the forces that drive cultural accumulation would help to explain the emergence of behavioral modernity, the changing rate of cultural evolution over time, and the variability in toolkit complexity among modern human groups.

The relationship between demography—in particular, population size and density—and cultural accumulation has been the focus of much attention, with construction and analysis of mathematical models (Aoki 2015; Aoki et al. 2011; Gilpin et al. 2016; Henrich 2004; Hochberg 2004; Kobayashi and Aoki 2012; Lehmann et al. 2011; Powell et al. 2009; Shennan 2001; Strimling et al. 2009), empirical studies (Caldwell and Millen 2010; Derex et al. 2013; Kempe and Mesoudi 2014; Muthukrishna et al. 2013), and statistical analyses of cultural complexity (Collard et al. 2005, 2013; Kline and Boyd 2010; Marquet et al. 2012; Read 2006). Influential theoretical treatments (Henrich 2004; Powell et al. 2009; Shennan 2001) and most laboratory experiments have concluded that demographic factors, especially population size, should predict cultural complexity. However, these findings appear to conflict with many statistical analyses of toolkit complexity in ethnographically described hunter-gatherer populations (Collard et al. 2005, 2013, 2016; Read 2006), which suggest that proxy measures of risk, such as above-ground productivity and mean effective temperature, not demographic factors, predict cultural complexity. On the other hand, an analysis by Kline and Boyd (2010) of the food-getting technology of relatively isolated and ecologically similar Oceanic island populations demonstrated significant relationships between both toolkit size and toolkit complexity and population size, but not environmental factors. These results, taken together with those of Collard et al. (2005, 2013, 2016), may suggest that risk drives cultural accumulation, but that population size becomes more important when environmental variation is minimal.

Many models do not investigate or test this relationship. Most models hold the innovation rate constant in order to assess the effects of population size (here and throughout the manuscript, following Cavalli-Sforza and Feldman 1981, “innovation” refers to the origination of new traits). Models of the evolution of genetic mutation rate (the genetic analogue of innovation rate) show that an increase in mutation rate is adaptive in an unpredictable environment (Carja et al. 2014; Gaál et al. 2010; Kussell and Leibler 2005; Lachmann and Jablonka 1996; Liberman et al. 2011). There may be a similar relationship between environmental instability and cultural innovation, suggesting that “necessity is the mother of invention.” This idea is supported by work on toolkit complexity (Collard et al. 2013), but it is sometimes disputed (Diamond 1998). Archaeological analyses over long time scales lend support to the idea that innovation is often linked to periods of environmental change or large migrations (Ziegler et al. 2013; but see Roberts et al. 2016).

To address the relationship between toolkit complexity, mode of cultural transmission, and population size, Aoki et al. (2011) introduced a class of stochastic population models from which they calculated the rate of cultural evolution under each of four modes of learning. These models and subsequent extensions were based on the Moran model of reproduction in a finite population and assumed an infinite-sites model of innovation (Ewens 2004:297). However, in reality, individuals have limited, not infinite, cognitive capacities, and a finite-sites model might be more appropriate. An increase in this cognitive limit may explain the rapid increase in cultural complexity in populations of modern humans compared with Neanderthals (Klein 2008; Klein and Steele 2013), although some argue that this does not account for the patchy presence of behavioral modernity in Middle Stone Age Africa prior to its more stable appearance in Upper Paleolithic Eurasia (Powell et al. 2009).

In order to explain the data on cultural complexity, a model should address the possible gain and loss of complexity over time as seen in present-day human populations; interactions with demography and environmental parameters; and the patchy presence of modern behavior in the archaeological record. Here, we suggest that the relationship between population size, cognitive capacity, and cultural complexity is more nuanced than previously appreciated. We show that if cognitive capacity is low relative to the population size, the rate of evolution and the eventual number of cultural traits in the population will be low compared with previous predictions based on infinite sites models (Aoki et al. 2011). We also show that when the cognitive capacity is high relative to the population size (such that the infinitely many sites model is relevant), the rate of evolution will be strongly influenced by the innovation rate, which may be correlated with both environmental unpredictability (Reader and Laland 2003) and population size. Therefore, in modern human populations we would expect to see a positive correlation either between population size and cultural complexity, where the environment is relatively benign and stable, or between risk and cultural complexity, where the environment is harsh and unpredictable.

## The Models

Here we present a combination of analytical and agent-based simulation models based on Aoki et al. (2011) and Fogarty et al. (2015b). We investigate the implications of limiting the possible number of cultural traits and extend the model to allow for increasing or decreasing population sizes to directly address the maladaptive loss of cultural traits (i.e., the “Tasmanian question”: see Diamond 1978; Henrich 2004; Jones 1977).

We explore the evolution of culture in which each individual is represented by a vector whose components are 0 or 1. Each vector is regarded as the cultural repertoire of an individual. We conduct agent-based simulations and focus on two questions. First, how are demography and cultural complexity related when we impose limits on individual cognition by making the vector finite? Second, what are the effects of population growth or decline, mode of transmission, rates of innovation, and the efficiency of social learning on the number of cultural traits at equilibrium? We adopt the following simulation framework.

Consider a population of  $N$  individuals in which individual  $l$  can be represented by a vector  $C_l$ , where the value for each trait  $i$  is denoted by  $c_{li}$ ; that is,  $C_l = \{c_{l,1}, c_{l,2}, \dots, c_{l,M}\}$ ,

where  $M$  is the total number of possible cultural traits. Thus, an individual is represented by a “trait vector” in which each trait takes a value of 1 or 0, representing presence or absence, respectively. We refer to  $M$  as the *imagination space*, which is considered to be under some sort of genetic control for the analyses presented here but could, with extensions to the model, change with the amount of culture already present (Fogarty et al. 2015a; Mesoudi 2011). Mesoudi (2011) introduced a limit on the cultural capacity of an individual, which he referred to as an “effort budget.” This budget allowed an individual to learn socially and innovate until his/her effort budget was exhausted. Here, the imagination space acts in a similar way, constraining the number of traits an individual can learn, and influencing the effective rate of innovation at the population level.

To begin our simulations, each of the  $N$  adult individuals is chosen in random order and given a chance to innovate. Each such individual may innovate a random trait that is monomorphic for 0 in the population in what we label the “infinite sites” innovation regime, or for which they have the 0 form, regardless of the values in the rest of the population, in what we call the “recurrent” innovation regime. The value of this trait in the individual’s trait vector changes from 0 to 1 with probability  $\mu/N$ . That is,  $\mu/N$  is the innovation rate per individual per time-step, where  $N$  time-steps constitute a generation and  $\mu$  is the rate per individual per generation. This process repeats until there are no remaining monomorphic traits (in the case of infinite sites innovation), or all individuals have had one chance to innovate.

After innovation, a naïve individual (with all traits of form 0) is born (making a total of  $N + 1$  individuals) and learns with probability  $\beta$  by choosing a role model in one of four different ways: (1) by *random oblique transmission*: the newborn chooses a role model randomly from the population of  $N$  adults and adopts its complete trait vector; (2) by *success-biased transmission*: the newborn samples  $k$  random individuals from the population of  $N$  adults, selects the individual with the most traits of form 1, and adopts its complete trait vector; (3) by *best-of- $k$  transmission*: the newborn samples  $k$  random individuals from the population. For each trait, if any of the newborn’s  $k$  role models have the 1 form of the trait, the newborn will copy that; if not, the newborn will retain 0. In this way the newborn pieces together its trait vector with input from up to  $k$  role models. And finally, (4) by *one-to-many transmission*: one individual is chosen at random and is the only available role model until its death. At birth, a newborn adopts that role model’s complete trait vector. In the following analyses we report the results for  $k = 2$  for the success-biased and *best-of- $k$*  transmission schemes. This value was chosen because there is a large difference in cultural accumulation between  $k = 1$  (or random oblique learning) and  $k = 2$ , but greater values of  $k$  do not have much incremental effect, as shown by Fogarty et al. (2015b).

In each of the four cases, the newborn replaces a random adult in the population. For *one-to-many transmission*, if the individual replaced is the role model, a new role model is selected at random from the population of  $N$  adults (including the latest newborn). If the individual replaced is not the role model, the same role model is retained for the next time-step. The simulation repeats this process of innovation → birth → learning → death  $N$  times until a new generation of  $N$  individuals is constituted. The whole process is repeated for a specified number of generations.

To calculate the rate of evolution for comparison with previous work (e.g., Aoki et al. 2011), we set the initial form of each trait to 0 in each individual at the outset of the simulation. Once a trait reaches fixation (i.e., all  $N$  individuals have the 1 form of that trait), that site is reset to 0 in all individuals, allowing further innovation at the site. The number

of fixation/reset events is counted and divided by the number of generations in the simulation run to give the rate of evolution. This is directly comparable to the analytical calculations of Aoki et al. (2011) for the infinite cultural sites model.

The parameter  $\beta$  represents the probability of transmission of the 1 form of each cultural trait (Lehmann et al. 2011).  $\beta$  is a per-trait probability of copying, and traits are copied independently. When computing the rate of evolution we set  $\beta = 1$ . If copying fails (probability  $1 - \beta$ ), the newborn retains the 0 form of the cultural trait. Each simulation is run for  $T = (nGen)N$  time steps where  $nGen$  represents the number of generations. The mean number of cultural traits in the population and other measures of cultural complexity are then calculated over the second half of each simulation run, from generation  $nGen/2$  on. Analytical expressions for the accumulation of traits for *random oblique* and *best-of-k* learning based on models described in Strimling et al. (2009) and Fogarty et al. (2015b), which assume  $M = \infty$ , are presented in ESM §1.

## Population Growth and Decline

Both the simulation models described above and their analytical analogues (Aoki et al. 2011; Strimling et al. 2009; see ESM §1) have assumed a constant population size. It is of course possible to compare populations of different constant sizes and infer, from this comparison, the relationship between population growth or decline and cultural complexity. However, this type of comparison neglects the important role of historical contingency. To address this issue directly, we use the modelling framework described in the previous section but introduce an analytical formulation that allows the population to grow or contract over time such that after every  $S$  time steps the population undergoes a birth-only event (in the case of population growth) or a death-only event (in the case of population decline). This means that the population size changes by one individual every  $S$  time steps. As shown in ESM §2, this formulation entails exponential change in population size.

We focus on the average number of cultural traits in a population (over individuals;  $\bar{C}_{pop}$ ) as a measure of cultural complexity and denote by  $\bar{C}_{pop,t}$  the average number of cultural traits present in the population at time  $t$ , measured in time steps.  $\bar{C}_{pop,t} = \sum_{i=1}^{N_t} P_{i,t}$ , where the subscript  $t$  denotes time and  $P_{i,t}$  is the expected number of cultural traits for which the 1-variant occurs in  $i$  individuals at the beginning of time step  $t$  (see ESM §2). We explore the temporal dynamics of  $\bar{C}_{pop,t}$  through numerical iteration with either a growing or a declining population as outlined above.

## Results

### Accumulation of Cultural Traits

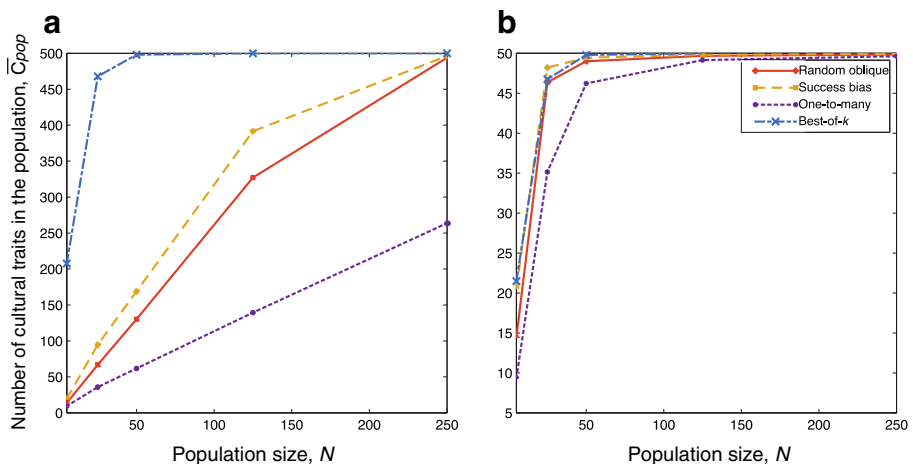
The simulated rate of cultural evolution matches the analytical values for the rate of evolution obtained in Aoki et al. (2011) when  $M$  is high relative to  $N$  (see the figure in ESM §1). In our fixed-population-size simulations, the number of cultural traits at

equilibrium increases with the size of the population, as shown in previous analyses (Aoki et al. 2011; Henrich 2004; Kempe et al. 2014; Lehmann et al. 2011; Shennan 2001; Strimling et al. 2009). However, both the rate of evolution and the number of cultural traits in a population exhibit a kind of saturation when  $M$  is low or the innovation rate,  $\mu$ , is high. This saturation effect occurs when there are relatively few sites available for innovation, with the consequence that the effective rate of innovation in the simulations is lower than the nominal rate (data not shown).

Fig. 1 shows that the number of cultural traits in the population has a roughly linear increase with population size before saturation, but the number of traits slowly reaches a plateau at  $M$  as  $N$  increases. This result is seen for both infinite sites mutation (shown in Fig. 1.) and for recurrent mutation (not shown) and draws attention to an intuitive but important point relevant to comparisons of the cultural evolution of earlier hominid species and our own: we expect the relationship between population size and cultural accumulation to be strong when imagination spaces are large, as is likely for modern humans, but population sizes are small. However, we expect a weak relationship or no relationship between population size and cultural accumulation for large populations with relatively smaller imagination spaces. We return to this point in the discussion.

## Cultural Diversity

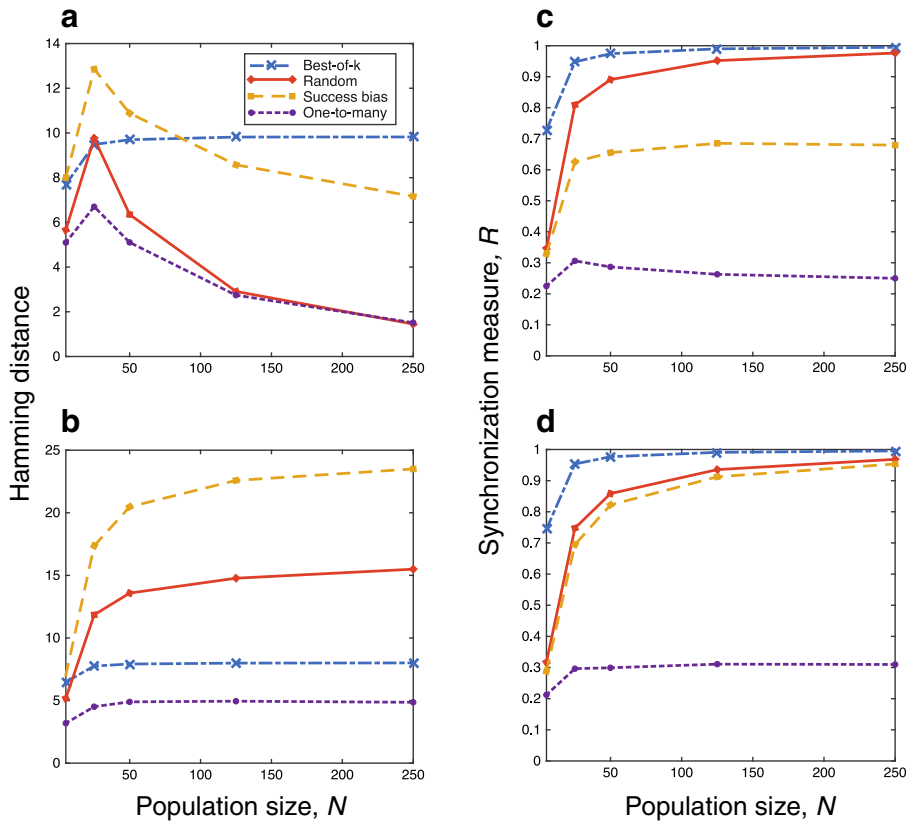
Another important metric of cultural complexity is cultural diversity. This can be quantified by calculating the average number of cultural traits at which two random individuals differ, namely the pairwise Hamming distance. The Hamming distance is a simple count of the entries in the trait vectors that differ between two individuals. For example, with  $M = 4$ , if individual A has trait vector  $\{1,1,0,0\}$  and individual B has trait vector  $\{0,1,0,1\}$ , the differences between the vectors occur at the first and last entries and the Hamming distance



**Fig. 1** The average number of cultural traits in the population,  $\bar{C}_{pop}$ , over the final 1000 simulated generations with infinite-sites mutation where the number of generations,  $nGen$ , was 2000,  $N = 5, 25, 125, 250$ , with  $\mu = 1$  (i.e., innovations arise once per individual per generation),  $\beta = 0.9$ ,  $k = 2$  where relevant. For panel a,  $M = 500$ , and for b,  $M = 50$

is 2. The average of the quantity is taken over all  $[N(N - 1)]/2$  pairs in the population and can be regarded as a population-level measure of cultural diversity.

The average Hamming distance for each mode of social learning in the infinite sites innovation model increases and then sharply decreases as a function of population size (Fig. 2a) when  $M = 50$  but differs strikingly in the model allowing recurrent innovation (Fig. 2c). This is due to the fundamentally different roles of innovation and social learning. The effect of innovation is to increase cultural distance between individuals—each innovation occurs only once in the infinite sites model and thus increases the distance between the innovating adult and the rest of the population. Conversely, the effect of social learning is to spread traits and to decrease the distance between individuals as they learn from one another. In the infinite sites regime, each innovation introduces a new trait to the population and simultaneously reduces the number of monomorphic 0 sites available for change, reducing the effective innovation rate. When the innovation rate is low, the effect of social learning is stronger than the effect of innovation on the distance between individuals, causing a sharp decrease in the average Hamming distance between individuals in the population. This is not



**Fig. 2** Measures of cultural diversity. Panels a and c show the average pairwise Hamming distance in the population, and b and d show the synchronization measure,  $R$ , over the final 1000 generations for  $N = 5, 25, 125, 250$ , with  $\mu = 1$  (i.e., innovations arise once per individual per generation),  $\beta = 0.9, k = 2$  where relevant, and  $M = 50$ . Panels a and b show the infinite-sites mutation model and c and d show results for the recurrent mutation model



seen with recurrent innovation and is strongest when the “saturation effect” of the infinite sites assumption becomes important—in other words, when  $M$  is low relative to  $N$ —but can also be seen to a lesser extent when  $M$  is higher. Note that this pattern does not hold for *best-of- $k$*  learning, in which case saturation occurs very rapidly.

However, because 0 represents the default trait of a newborn individual, another measure may be needed to assess whether similarity within populations is, in fact, due to cultural transmission and not simply to this default. To measure a synchronization effect, we calculate the ratio,  $R$ , of observed Hamming distance to its expected value under random matching. This is given by

$$R = (\text{Hamming distance})/2Mp(1-p)$$

where  $p = \frac{\bar{C}_{ind}}{M}$  is the frequency of the 1 variant (with  $\bar{C}_{ind}$  representing the average number of cultural traits of form 1 per individual in the population) and  $2p(1-p)$  is the expected value of the Hamming distance per trait when  $\bar{C}_{ind}$  variants of value 1 are randomly distributed among  $M$  traits for two individuals.  $R$  is smaller when the synchronization effect is stronger. *One-to-many* transmission has the strongest synchronization effect since all newborns copy the whole sequence of one role model. *Success-biased* transmission also shows strong synchronization among individuals since a smaller number of successful individuals are preferred as role models. Finally, *best-of- $k$*  transmission has a weak synchronization effect because pupils copy individual traits rather than whole trait vectors (Fig. 2b and d).

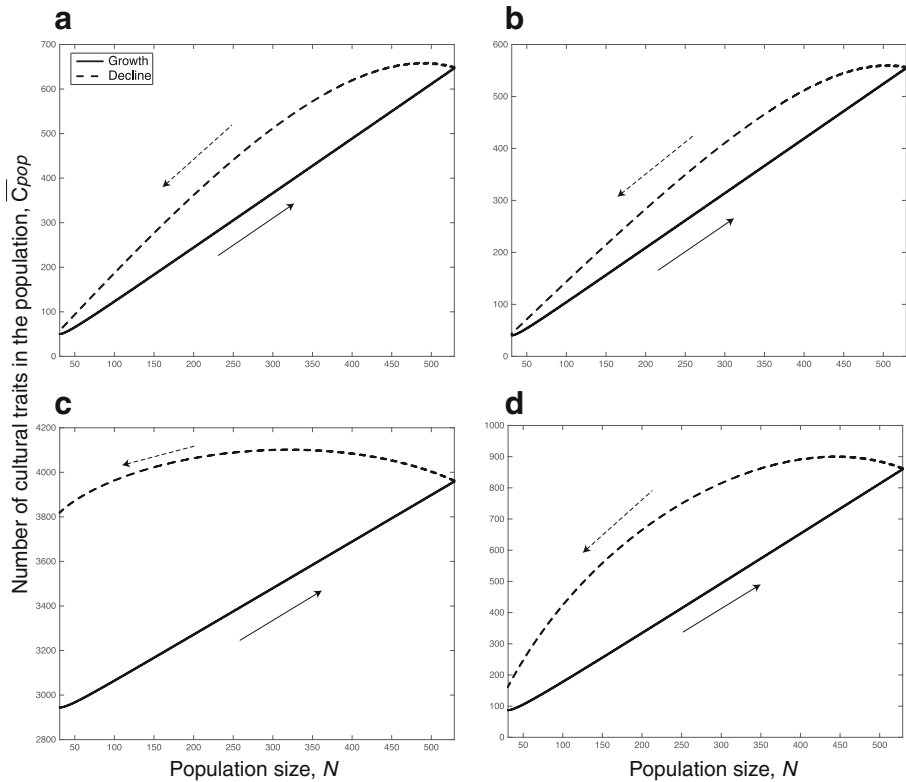
## Population Growth and Decline

Finally, we examine the important cases of population growth and decline. Here, we separate the effects of population size and population growth and investigate whether modes and fidelity of transmission affect the accumulation of traits as population size grows or falls. The panels in Fig. 3 each show a population of a size consistent with a hunter-gatherer band ( $N = 30$ ), growing to a size consistent with a hunter-gatherer tribe ( $N = 530$ ), and an initially large population ( $N = 530$ ) in decline. For *random oblique* transmission, panels 3a and 3b show the number of cultural traits in the population over time as the small population grows and as the large population shrinks, and panels 3c and 3d show the same for *best-of- $k$*  transmission. The learning fidelity is relatively high ( $\beta = 0.7$ ) in panels 3a and 3c and relatively low ( $\beta = 0.5$ ) in panels 3b and 3d. The increase in  $\bar{C}_{pop}$  during population growth is approximately linear (see ESM 2).

The population is allowed to reach equilibrium in  $\bar{C}_{pop}$  at  $N = 30$  before growing to  $N = 530$ . Population decline then begins as soon as the population reaches 530 individuals, and  $\bar{C}_{pop}$  can be assessed during both growth and decline. We note that the results are qualitatively similar when we allow populations to reach equilibrium at the larger population size before this size begins to decline. However, in most cases reaching this equilibrium takes a considerable number of generations, and the magnitude of  $\bar{C}_{pop}$  at these equilibria makes this an unlikely scenario in a natural setting.

In the case of *random oblique* transmission, in a growing population cultural traits tend to accumulate, whereas in a declining population the number of cultural traits tends to decline. This suggests that population decline can cause substantial cultural loss in the case of *random oblique* transmission even for easy-to-learn traits ( $\beta = 0.7$ , Fig. 3a). However, Fig. 3c shows that this is not the case for all modes of transmission. With the “faster” *best-of- $k$*  transmission





**Fig. 3** The average number of cultural traits ( $\bar{C}_{pop}$ ) in a population undergoing growth from the size of a small human band ( $N = 30$ ) to a larger population ( $N = 530$ ) (solid) and decline from  $N = 530$  to  $N = 30$  (dashed) for *random oblique* (a and b) and *best-of-k* (c and d) transmission modes calculated using eq. S2.1 in the ESM.  $\beta = 0.7$  for panels a and c and  $\beta = 0.5$  for panels b and d. In panels c and d,  $k = 2$ . In all panels,  $S = 5$ . Arrows indicate the direction of time. Note the differences in y-axis scales

(the *best-of-k* model predicts a higher rate of transmission than random oblique), as population size declines, the rate of accumulation of traits slows but the number of cultural traits can continue to increase for easy-to-learn traits ( $\beta = 0.7$ ) and, importantly, the value of  $\bar{C}_{pop}$  when  $N$  reaches 30 is considerably higher than the equilibrium for that value of  $N$ . For difficult-to-learn traits ( $\beta = 0.5$ ), panel 3d shows that a decline in population size can cause loss of cultural traits even with the faster *best-of-k* learning. This supports the findings of previous analyses, which suggested that more complex or difficult-to-learn traits are lost more easily than traits that are easy to learn (Fogarty et al. 2015b; Henrich 2004), with the caveat that this effect is more pronounced under faster cultural transmission, e.g., *best-of-k*. Figure 3 also shows that the number of cultural traits in a population depends not only on population size but also on the history of population growth or decline and the mode of cultural transmission.

### Discussion

In order to explain the divergence between standard theoretical expectations and ethnographic data on cultural complexity, we focused on cases where simulations did

not show a straightforward increase in cultural complexity with population size. We see this when  $M$  is low relative to  $N$ , meaning that the imagination space (or cognitive capacity) of the population is limited relative to the number of individuals in the population, in which case there is cultural saturation; that is, the availability of sites for innovation decreases sharply, causing the effective rate of innovation in the simulations to fall. The recurrent innovation model mitigates this effect by making non-monomorphic sites available for innovation. However, both innovation models show decreases in the cultural evolutionary rate and in the number of traits expected in the population compared with analytical predictions for the infinitely many sites model (Aoki et al. 2011) when  $M$  is low compared to  $N$ . This implies that the relationship between population size and cultural accumulation depends on complex interactions between innovation rate, cognitive capacity, and population size.

The lithic technology's slow rate of change between 600 and 100 kya (Klein 2008) along with its low diversity in the Lower and perhaps Middle Paleolithic (Stiner and Kuhn 2006) may be explained if ancestral hominins had both a limited cognitive capacity (low  $M$ ) relative to modern humans and a sufficiently large population size (comparable  $N$ ). This combination would result in a slow rate of cultural evolution (or evolutionary stasis), where  $\bar{C}_{pop}$  hardly increases (or does not increase) with increasing  $N$ . Powell et al. (2009) used an estimate of effective population size from coalescent studies of human mtDNA (Atkinson et al. 2008) to approximate population size for modern humans in the Upper Paleolithic, and some estimates of effective population size of populations ancestral to both Neanderthals and humans are becoming available (e.g. Fabre et al. 2009). Of course, effective population size may be a poor proxy for census population size, which is of greater relevance to questions of cultural evolution.

Some genetic evidence suggests that Neanderthal effective population size was similar to that of contemporaneous modern humans, although much smaller than that of present-day humans (Lalueza-Fox et al. 2011). For example, Atkinson et al. (2008) estimate the effective population size of modern humans in Europe at ~45kya to have been around 2900 and Briggs et al. (2009) estimate that the effective population size of Neanderthals remained small between 70 kya and 30 kya but was similar to that of modern humans; their estimate is 1470.

The cognitive capacity of Neanderthals is difficult if not impossible to assess accurately, and some recent discoveries suggest that they were capable of at least rudimentary engravings (Rodríguez-Vidal et al. 2014). However, Mithen (1996) suggests that Neanderthals might have been unable to understand certain types of information in the same way as modern humans. For example, he suggests that they might have been unable to understand complex cross-domain constructs such as metaphors or symbolic representation. If this were the case, the number of cultural traits Neanderthals were capable of producing would have been limited compared with that of modern humans, who did not share these cognitive restrictions. Overall, given parameters that reflect this scenario (low  $M$  and relatively high  $N$ ), our model predicts that Neanderthals should have had a slow or zero rate of cultural evolution as their population size grew, after an initial burst of relatively rapid cultural evolution when their population size was small.

The opposite was true in the Upper Paleolithic for anatomically modern humans, who arguably had the large and flexible cognitive capacity that enables cross-domain

and metaphorical thinking in today's humans. As the former reached higher population sizes, their rate of cultural evolution continued to increase, leading to the explosion in cultural evolution that heralded behavioral modernity. This echoes the conclusion of Powell et al. (2009), who showed that behavioral modernity arrived in European human populations as population sizes grew on that continent.

We can also address the question of what predicts toolkit complexity in modern hunter-gatherer groups. Here, the cognitive capacity,  $M$ , is high, but the population size of such groups may be quite low, and the rate of cultural evolution and the amount of accumulation will be dominated by the rate of innovation or the population size rather than the interaction between cognition and population size, as suggested above for Neanderthals. If innovation rate is a major determinant of the rate of evolution in small groups of modern humans, we might expect to see correlations among the rate of evolution, cultural accumulation, and environmental factors affecting innovation, both historical and recent, and a weaker correlation with demographic factors except where the environment is similar for all populations, as in Kline and Boyd (2010).

This may provide a clue to the possible course of events in Tasmania, where seemingly adaptive cultural traits were lost when the Tasmanian Aboriginals were cut off from mainland Australia at the start of the Holocene. Read (2006) suggested that this loss of cultural complexity was an adaptive response to environmental pressures, and our model tentatively suggests that the low population size and high cognitive capacity of the Tasmanians would cause the rate of cultural evolution in terms of either loss or accumulation of traits to depend on environmental factors affecting innovation. Our numerical simulations further suggest that the decline in population size, not the low population size per se, could have caused rapid cultural loss (Fig. 3). This draws attention to one important, albeit difficult to model, aspect of this debate that has been neglected, namely, the importance of non-equilibrium situations and historical contingency. Previous mathematical models have assumed both a static population size and equilibrium levels of culture.

Further, we speculate that modes of cultural transmission that support rapid cultural evolution (such as *best-of-k* transmission) were less likely to have been a strong feature of Tasmanian learning, as these would have mitigated the loss of cultural traits in a way that other mechanisms (such as random oblique learning) might not. Evidence for the use of either of these transmission modes in any human population is scarce, but Henrich and Broesch (2011) suggest that other, more complex transmission rules, such as prestige bias, may be widely used.

Finally, it is interesting to consider the relevance of our model to recent experimental micro-society studies on cultural accumulation (Derex et al. 2013; Muthukrishna et al. 2013), which showed that the generation and maintenance of cultural diversity within the micro-societies depended on population size. The problem addressed by these studies differs in scale and scope from our simulations and population-wide statistical analyses. Although the experimenters considered just a single task, the number of possible innovations a participant could generate within the confines of that task was probably quite high. The population sizes, however, were relatively small (up to 16 in Derex et al. 2013, for example);

hence, the diversity within the set of task solutions increased with population size, as in our simulations when  $N$  is low relative to  $M$ .

Our analysis suggests that previous theoretical and empirical findings showing a strong relationship between cultural accumulation and demographic factors such as population size are not, as previously thought, incompatible with the findings from statistical analysis of cultural data, which indicate that cultural accumulation in modern hunter-gatherers depends on environmental factors such as above-ground productivity. A full understanding of the emergence of behavioral modernity, and the cultural evolution that has followed, depends on understanding and untangling the complex relationships between culture, cognition, and demography.

**Acknowledgments** The authors would like to thank Alex Mesoudi and two anonymous referees for their helpful reviews of this paper. The work was supported by a Templeton Foundation grant to MWF, Monbukagakusho grant 22101004 to KA, Monbukagakusho grant 16H06412 to JYW, and a 2020 Science Fellowship to LF at UCL.

## References

- Aoki, K. (2015). Modeling abrupt cultural regime shifts during the Paleolithic and stone age. *Theoretical Population Biology*, *100*, 6–12.
- Aoki, K., Lehmann, L., & Feldman, M. W. (2011). Rates of cultural change and patterns of cultural accumulation in stochastic models of social transmission. *Theoretical Population Biology*, *79*(4), 192–202.
- Atkinson, Q. D., Gray, R. D., & Drummond, A. J. (2008). mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. *Molecular Biology and Evolution*, *25*(2), 468–474.
- Briggs, A. W., Good, J. M., Green, R. E., Krause, J., Maricic, T., Stenzel, U., et al. (2009). Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science*, *325*(5938), 318–321.
- Caldwell, C. A., & Millen, A. E. (2010). Conservatism in laboratory microsocieties: unpredictable payoffs accentuate group-specific traditions. *Evolution and Human Behavior*, *31*(2), 123–130.
- Carja, O., Liberman, U., & Feldman, M. W. (2014). Evolution in changing environments: modifiers of mutation, recombination, and migration. *Proceedings of the National Academy of Sciences*, *111*(50), 17935–17940.
- Cavalli-Sforza, L., & Feldman, M. W. (1981). *Cultural transmission and evolution*. Princeton, NJ: Princeton University Press.
- Collard, M., Kemery, M., & Banks, S. (2005). Causes of toolkit variation among hunter-gatherers: a test of four competing hypotheses. *Canadian Journal of Archaeology*, *29*(1), 1–19.
- Collard, M., Buchanan, B., & O'Brien, M. J. (2013). Population size as an explanation for patterns in the Paleolithic archaeological record. *Current Anthropology*, *54*(S8), S388–S396.
- Collard, M., Vaesen, K., Cosgrove, R. & Roebroeks, W. (2016). The empirical case against the “demographic turn” in Palaeolithic archaeology. *Philosophical Transactions of the Royal Society of London, B: Biological Sciences*, *371*, 20150242. doi:10.1098/rstb.2015.0242
- Derex, M., Beugin, M.-P., Godelle, B., & Raymond, M. (2013). Experimental evidence for the influence of group size on cultural complexity. *Nature*, *503*(7476), 389–391.
- Diamond, J. (1978). The Tasmanians: the longest isolation, the simplest technology. *Nature*, *273*, 185–186.
- Diamond, J. M. (1998). *Guns, germs and steel: a short history of everybody for the last 13,000 years*. London: Random House.
- Ewens, W. J. (2004). *Mathematical population genetics: Theoretical introduction*. Interdisciplinary Applied Mathematics 27. New York: Springer-Verlag.
- Fabre, V., Condemni, S., & Degioanni, A. (2009). Genetic evidence of geographical groups among Neanderthals. *PLoS One*, *4*(4). doi:10.1371/journal.pone.0005151.
- Fogarty, L., Creanza, N., & Feldman, M. W. (2015a). Cultural evolutionary perspectives on creativity and human innovation. *Trends in Ecology & Evolution*, *30*(12), 736–754.

- Fogarty, L., Wakano, J. Y., Feldman, M. W., & Aoki, K. (2015b). Factors limiting the number of independent cultural traits that can be maintained in a population. In K. Aoki & A. Mesoudi (Eds.), *Learning strategies and cultural evolution during the Paleolithic* (pp. 9–21). Tokyo: Springer Japan
- Gaál, B., Pitchford, J. W., & Wood, J. A. (2010). Exact results for the evolution of stochastic switching in variable asymmetric environments. *Genetics*, *184*(4), 1113–1119.
- Gilpin, W., Feldman, M. W., & Aoki, K. (2016). An ecocultural model predicts Neanderthal extinction through competition with modern humans. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 2134–2139.
- Henrich, J. (2004). Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses: the Tasmanian case. *American Antiquity*, *69*(2), 197–214.
- Henrich, J., & Broesch, J. (2011). On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Philosophical Transactions of the Royal Society of London, B: Biological Sciences*, *366*(1567), 1139–1148.
- Hochberg, M. E. (2004). A theory of modern cultural shifts and meltdowns. *Proceedings of the Royal Society of London, B: Biological Sciences*, *271*, S313–S316.
- Jones, R. (1977). The Tasmanian paradox. In R. V. S. Wright (Ed.), *Stone tools as cultural markers: Change, evolution and complexity* (pp. 189–204). Canberra: Australian Institute of Aboriginal Studies.
- Kempe, M., & Mesoudi, A. (2014). An experimental demonstration of the effect of group size on cultural accumulation. *Evolution and Human Behavior*, *35*(4), 285–290.
- Kempe, M., Lycett, S. J., & Mesoudi, A. (2014). From cultural traditions to cumulative culture: parameterizing the differences between human and nonhuman culture. *Journal of Theoretical Biology*, *359*, 29–36.
- Klein, R. G. (2000). Archeology and the evolution of human behavior. *Evolutionary Anthropology*, *9*(1), 17–36.
- Klein, R. G. (2008). Out of Africa and the evolution of human behavior. *Evolutionary Anthropology*, *17*(6), 267–281.
- Klein, R. G., & Steele, T. E. (2013). Archaeological shellfish size and later human evolution in Africa. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(27), 10910–10915.
- Kline, M. A., & Boyd, R. (2010). Population size predicts technological complexity in Oceania. *Proceedings of the Royal Society of London, B: Biological Sciences*, *277*(1693), 2559–2564.
- Kobayashi, Y., & Aoki, K. (2012). Innovativeness, population size and cumulative cultural evolution. *Theoretical Population Biology*, *82*(1), 38–47.
- Kussell, E., & Leibler, S. (2005). Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, *309*(5743), 2075–2078.
- Lachmann, M., & Jablonka, E. (1996). The inheritance of phenotypes: an adaptation to fluctuating environments. *Journal of Theoretical Biology*, *181*(1), 1–9.
- Lalueza-Fox, C., Rosas, A., Estalrich, A., Gigli, E., Campos, P. F., García-Tabernero, A., et al. (2011). Genetic evidence for patrilocality among Neanderthal groups. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(1), 250–253.
- Lehmann, L., Aoki, K., & Feldman, M. W. (2011). On the number of independent cultural traits carried by individuals and populations. *Philosophical Transactions of the Royal Society of London, B: Biological Sciences*, *366*(1563), 424–435.
- Lieberman, U., Van Cleve, J., & Feldman, M. W. (2011). On the evolution of mutation in changing environments: recombination and phenotypic switching. *Genetics*, *187*(3), 837–851.
- Marquet, P. A., Santoro, C. M., Latorre, C., Standen, V. G., Abades, S. R., Rivadeneira, M. M., et al. (2012). Emergence of social complexity among coastal hunter-gatherers in the Atacama Desert of northern Chile. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(37), 14754–14760.
- Mesoudi, A. (2011). Variable cultural acquisition costs constrain cumulative cultural evolution. *PloS One*, *6*(3), 15–17. doi:10.1371/journal.pone.0018239.
- Mithen, S. (1996). *The prehistory of the mind*. London: Thames and Hudson.
- Muthukrishna, M., Shulman, B. W., Vasilescu, V., & Henrich, J. (2013). Sociality influences cultural complexity. *Proceedings of the Royal Society of London, B: Biological Sciences*, *281*(20132511).
- Powell, A., Shennan, S., & Thomas, M. G. (2009). Late Pleistocene demography and the appearance of modern human behavior. *Science*, *324*(5932), 1298–1301.
- Read, D. (2006). Tasmanian knowledge and skill: maladaptive imitation or adequate technology? *American Antiquity*, *71*(1), 164–184.
- Reader, S. M., & Laland, K. N. (2003). *Animal innovation*. Oxford: Oxford University Press.

- Roberts, P., Henshilwood, C. S., Van Niekerk, K. L., Keene, P., Gledhill, A., Reynard, J., Badenhorst, S. & Lee-Thorp, J. (2016). Climate, environment and early human innovation: stable isotope and faunal proxy evidence from archaeological sites (98–59ka) in the southern Cape, South Africa. *PLoS One*, *11*, 1–20.
- Rodríguez-Vidal, J., d’Errico, F., Pacheco, F. G., Blasco, R., Rosell, J., Jennings, R. P., et al. (2014). A rock engraving made by Neanderthals in Gibraltar. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(37), 13301–13306.
- Shennan, S. (2001). Demography and cultural innovation: a model and its implications for the emergence of modern human culture. *Cambridge Archaeological Journal*, *11*(01), 5–16.
- Stiner, M. C., & Kuhn, S. L. (2006). Changes in the “connectedness” and resilience of Paleolithic societies in Mediterranean ecosystems. *Human Ecology*, *34*(5), 693–712.
- Strimling, P., Sjöstrand, J., Enquist, M., & Eriksson, K. (2009). Accumulation of independent cultural traits. *Theoretical Population Biology*, *76*(2), 77–83.
- Ziegler, M., Simon, M. H., Hall, I. R., Barker, S., Stringer, C., & Zahn, R. (2013). Development of Middle Stone Age innovation linked to rapid climate change. *Nature Communications*, *4*(May), 1905. doi:10.1038/ncomms2897.

**Laurel Fogarty** is a postdoctoral 2020 Science research fellow working at CoMPLEX in University College London. She conducts research on mathematical and computational models of cultural evolution and the evolution of communication.

**Joe Yuichiro Wakano** has been an associate professor at Meiji University since 2007. He conducts research on mathematical biology and mathematical analyses on evolutionary models, including cultural evolution.

**Marcus Feldman** is the Burnet C. and Mildred Finley Wohlford Professor of Biological Sciences and director of the Morrison Institute for Population and Resource Studies at Stanford University. He uses applied mathematics and computer modeling to simulate and analyze the processes of genetic and cultural evolution.

**Kenichi Aoki** is an emeritus professor at the University of Tokyo, where he taught anthropology before retiring in 2012, and is currently visiting research associate at Meiji University, where he conducts research on mathematical biology.