ABSTRACT  The accumulated copying error (ACE) model exploits the fact that limits to human visual perception introduce proportional copying errors during cultural transmission of a continuous trait, such as the length of a tool. Archaeologists have used ACE to infer the mechanisms of cultural transmission employed in the past. But ACE’s predictions apply only under “idealized” conditions in which the population contains infinitely many vertical transmission chains and the continuous trait can take any positive value. Here, I relax each assumption to show: (1) the mean and variance of a vertically transmitted continuous trait are functions of population size and number of transmission events, and (2) functional constraints weaken the cumulative effects of proportional copying error. Incorporating the effects of population size and functional constraints in future work will reduce Type I errors (falsely rejecting a true null) when inferring the cultural transmission mechanisms responsible for population-level variation in continuous traits. [cultural evolution, cultural transmission, finite population size, functional constraints, Paleolithic, simulation, social learning]

RESUMEN  El modelo de error de copia acumulado (ACE) explota el hecho que límites a la percepción visual humana introducen errores de copia proporcionales durante la transmisión cultural de un rasgo continuo, tal como la longitud de una herramienta. Arqueólogos han usado ACE para inferir los mecanismos de transmisión cultural empleados en el pasado. Sin embargo, las predicciones del ACE aplican sólo bajo condiciones “idealizadas” en las cuales la población contiene infinitamente muchas cadenas de transmisión vertical y el continuo rasgo que puede tomar cualquier valor positivo. Aquí, flexibilizo cada asunción para mostrar que: (1) la media y varianza de un rasgo continuo transmitido verticalmente son funciones del tamaño de la población y el número de eventos de transmisión, y (2) limitaciones funcionales debilitan los efectos acumulativos del error de copia proporcional. Incorporar los efectos del tamaño de la población y las limitaciones funcionales en trabajo futuro reducirá los errores Tipo I (rechazar falsamente una hipótesis nula verdadera) cuando infiriendo los mecanismos de transmisión cultural responsables de la variación del nivel de población con rasgos continuos. [evolución cultural, transmisión cultural, tamaño de población finita, limitaciones funcionales, Paleolítico, simulación, aprendizaje social]
Cultural transmission is powerful but imperfect. We do not simply receive cultural information from others as if it were a package in the mail. Instead, we first perceive and then attempt to reproduce the information to which we have been exposed. We do not always get it exactly right; to err is human, after all. The accumulated copying error model (hereafter, ACE) cleverly exploits the fact that limitations to human visual perception inevitably introduce errors during the transmission of continuous cultural traits. A continuous trait (also referred to as a quantitative trait) is any metric attribute of material culture that can be measured on the continuous scale; examples include the length of a stone tool or the thickness of a ceramic vessel. And, as reviewed below, the fact that the size of the error introduced during cultural transmission is proportional to the target value of the continuous trait holds important implications for how the mean and variance of the trait ought to change through time in a constant-sized population (Eerkens and Lipo 2005; Hamilton and Buchanan 2009; Kempe, Lycett, and Mesoudi 2012).

ACE generates expectations against which researchers can compare empirical observations to help infer the mechanisms by which the continuous trait of interest was transmitted. Eerkens and Lipo (2005) compare ACE’s expectations to quantitative measures of Archaic projectile points and Woodland ceramics, Hamilton and Buchanan (2009) to those of Clovis projectile points, and Kempe, Lycett, and Mesoudi (2012) to those of experimental and archaeological Acheulean handaxes. These researchers interpret empirical departures from ACE’s expectations as potential evidence of biased cultural transmission, such as conformity or prestige bias, and/or of external selective pressures on tool design. But previous applications have not yet interrogated two of ACE’s fundamental, and potentially problematic, assumptions: (1) that the population contains infinitely many transmission chains; and (2) that the continuous trait of interest can take any positive value, no matter how large or small. Obviously, no empirical case includes infinitely many transmission chains, simply because no population is infinitely large. Nor can a continuously varying attribute (e.g., length, width, thickness) of any handmade implement become infinitely large or infinitesimally small, no matter how many times it is subject to proportional copying error. It is important that we understand the effects of violating each of these assumptions before comparing ACE’s expectations against empirical data. To the extent that finite population size and functional constraints dampen the cumulative effects of proportional copying error, they also weaken the inferential power of ACE’s predictions. The results of my study show why the expectations provided by a new formal model that accounts for finite population size and functional constraints are more appropriate, albeit ultimately no less fraught.

The article is structured as follows. First, I use my review of ACE as an opportunity to reconcile previous studies’ contradictory conclusions regarding the cumulative effect of proportional copying error on the “expected value”—i.e., the mean—and the variance of a vertically transmitted continuous cultural trait in an infinitely large population. After introducing a new version of ACE that disallows perceptible errors, I use simulation to show how sampling bias affects the mean and variance of a vertically transmitted continuous trait in any finite population. I then investigate whether population size has similar effects under unbiased, prestige biased, and conformist biased cultural transmission. Next, I show that a continuous trait’s functional constraints severely limit the cumulative effects of proportional copying error. I reassess ACE’s utility in light of my results before concluding with some thoughts on how future studies might build upon the foundation ACE provides to improve our ability to infer cultural transmission mechanisms from population-level data on continuous traits.

ACE AT FIFTEEN: A REVIEW AND RECONCILIATION

The notion that limitations to human perception introduce errors during the transmission of information will come as little surprise to those who have played the parlor game “broken telephone.” In broken telephone, players sit in a circle and whisper a message to their neighbor (for a useful review of broken telephone, see Mesoudi 2011, 139–41). Each whisper serves as a link in the “transmission chain” of players. Each transmission event provides an opportunity for error in enunciation on the part of the speaker and/or interpretation on the part of the listener. Even if everyone is playing in good faith, the original sentence changes as small errors get incorporated into the message as it moves from person to person through the chain. Due to the cumulative effects of auditory perception error, a game of broken telephone that is initialized with “Mike called for pizza with cheese and mushrooms” might end with the final person in the chain proclaiming he heard “My cousin lives near trees in Tucson.”

Clearly, humans are not broken telephones, and important technological information is not commonly transmitted by hushed whispers in the midst of a noisy party. In more mundane settings, information regarding the design of a tool is transmitted from a more experienced individual to a less experienced individual in such a way as to reduce the very errors encouraged by the rules of broken telephone. So then what level of perception error remains under more realistic conditions in which novice toolmakers are allowed to seek out clarification and additional direction from experts? How do the fundamental limits of our ability to perceive and reproduce a continuous attribute of an implement by hand in stone, clay, wood, or bone affect the mean and variance of that attribute through time? By inquiring as to the expected mean and variance of a continuous trait under the highest fidelity of transmission possible for humans, ACE models the “best-case scenario,” whereby other potential sources of error, such as variation in the quality of the raw materials used to make a tool, are purposefully excluded (but see Hamilton and Buchanan [2009] for an example that includes a term for “structural error”). The answer to this question ought to be useful for those who wish to infer the mecha-
nism(s) of cultural transmission responsible for the observed mean and variance of an empirical continuous trait of interest. In previous studies, empirical departures from ACE’s expectations—that is, observed values that are either larger or smaller than one’s expectations—have been interpreted as evidence that the trait was passed via a mechanism other than vertical cultural transmission.

The game of broken telephone exploits auditory perception error. But because ACE concerns continuous attributes of handmade objects, it addresses visual perception error. Interestingly, psychophysical studies suggest the size of the error introduced by visual perception is proportional to the “target value,” or the value of the continuous trait (e.g., length, width) one is trying to reproduce from memory. Previous work shows that humans generally have difficulty distinguishing between lines that differ in length by less than 3 percent unless given recourse to a ruler or the chance to place the lines directly side by side (Coren, Ward, and Enns 1994; Eerkens 2000; Eerkens and Bettinger 2001; Gilinsky 1951; Norwich 1987; Ono 1967), suggesting that the so-called Weber Fraction for visual perception limits the size of the error that is visually imperceptible to humans as a proportion of the target value. This finding has far-reaching implications for the study of cultural transmission with archaeological data. Eerkens and Lipo (2005) were the first to propose that folding the Weber Fraction for visual perception error into a broken-telephone-like model of cultural transmission provides a tool for estimating both the mean and the level of variation one would expect to see at a continuous trait in a population after a given number of vertical transmission events.

We can now begin to describe ACE more formally. For the purposes of this article, I use X to represent a quantifiable and continuously varying physical attribute (e.g., length) of some class of culture material (e.g., projectile point) and t to refer to the number of times this attribute has been transmitted from an “experienced” individual to a “naïve” individual within each of N transmission chains. In the real world, the actual duration of time—hours, days, months, years, etc.—represented by each transmission event, t, will vary among traits, with some traits transmitted more regularly (with a shorter periodicity) than others. In the interest of keeping the model general, I leave it to the reader to define the periodicity of t that is appropriate for the continuous trait one has in mind. At any rate, XN, t represents the set of continuous trait values displayed by a finite population of N (i.e., N < ∞) transmission chains after t transmission events in each chain. For clarity, I use X∞, t to designate the special case where N = ∞. Note that while XN, t refers to a set of continuous values, I use the lowercase xi,t to represent the value of the continuous trait displayed by the individual in the i-th chain after the t-th transmission event.

With proportional copying error at its heart, ACE describes a multiplicative stochastic process—a seemingly esoteric detail that nonetheless is crucial for understanding the population size effects discussed below. Multiplicative stochastic processes exhibit nonadditive properties that distinguish them from more intuitive additive stochastic processes. To illustrate this difference with a simple deterministic example, consider how much easier it is to calculate how much money one will save over the course of twenty years by socking away 100 dollars per month (100 dollars * 12 months * 20 years) than it is to calculate how much money will accumulate over the same period if one instead invests 100 dollars each month in a special savings account that accrues 5 percent interest per month (now wouldn’t that be grand?). The former is a strictly additive process by which one’s nest egg increases at a constant rate of 100 dollars per month. By contrast, the process describing the monthly rate of increase in the latter is partly multiplicative due to 5 percent monthly interest. Although in both scenarios the same amount of money is deposited each month, only in the latter does the savings increase at an increasing monthly rate: in the first month the amount increases by (100 + (0 * 0.05)) = $100, in the second month it increases by (100 + (100 * 0.05)) = $105, in the third month it increases by (100 + (205 * 0.05)) = $110.25, and so on. Because it is difficult to account for savings that increase at an increasing rate, people intuitively underestimate the final balance in the multiplicative scenario ($243,477,147.48) but not the additive scenario ($24,000).

Due to the effects of visual perception error, whereby the size of the error introduced during each transmission event is proportional to—and thus dependent upon—the target value, an infinitely large population of geometric Brownian random walks yields a log-normally distributed random variable (called X∞, t), bound by 0 on the left and positively skewed to the right. Although the expected value—i.e., the mean—of this log-normally distributed random variable does not change with increasing t (Gardiner 2004, 103–4; Kempe, Lycett, and Mesoudi 2012), its variance increases exponentially with increasing t (Kempe, Lycett, and Mesoudi 2012; Limpert, Stahel, and Abbt 2001). Just as one’s savings increases at an increasing rate due to compounding 5 percent monthly interest, the distribution of X∞, t becomes increasingly right-skewed at an increasing rate with increasing t thanks to compounding proportional copying error. As Kempe, Lycett, and Mesoudi (2012) correctly point out, for sufficiently large t, ACE predicts that while the vast majority of transmission chains in an infinitely large population will display values much smaller than the value they started with at t = 0, a very small proportion will display values that are many orders of magnitude greater than their values at t = 0.

ACE assumes that the continuous trait of interest is transmitted vertically from the “experienced generation” t to the “naïve generation” t + 1. Vertical cultural transmission describes the case in which each “experienced” practitioner transmits information to one and only one novice, or “naïve,” individual. Assuming that N remains constant, vertical transmission holds that each novice learns from a different member of the experienced generation. In the context of
cultural evolution, vertical transmission does not require that the experienced practitioner is a relative, let alone the biological parent, of the novice. Within each transmission chain, \( x_{i,t+1} \) is a function of the value displayed by the naïve individual’s teacher (i.e., \( x_{i,t} \)) plus the proportional error introduced during cultural transmission. Despite all of the mathematical notation and jargon, ACE merely represents a normally distributed error term. Because Eq. 1 describes a multiplicative stochastic process, which as explained above results in a log normally distributed random variable (Hamilton and Buchanan 2009; Limpert, Stahel, and Abbt 2001). The conceptual model underlying ACE holds that “just noticeable differences”—errors with an absolute size greater than \((w \times x_{i,t})\)—would not be tolerated or passed on. However, as a consequence of employing the normal distribution with mean 0 and variance \((w/2)^2\), approximately 4.55 percent of transmission events result in copying errors with absolute values that surpass \((w \times x_{i,t})\). Also problematic is the fact that drawing a value from the normal distribution makes it possible—although highly unlikely, given a small \(w\)—for \( x_{i,t+1} \) to take a negative value even though there is no such thing as a negative length, width, or thickness.

Hamilton and Buchanan (2009, Eq. 2) and Kempe, Lycett, and Mesoudi (2012) represent the same multiplicative random process with a slightly different time-series equation:

\[
x_{i,t+1} = x_{i,t} \times \left[ 1 + \mathcal{N}\left(0, \left(\frac{w^2}{2}\right)^2\right)\right],
\]

where \( w \) is the Weber Fraction and \( \mathcal{N}(0, (w/2)^2) \) is a normal random variable with mean 0 and variance \((w/2)^2\). Eq. 1 describes a multiplicative stochastic process, which as explained above results in a log normally distributed random variable (Hamilton and Buchanan 2009; Limpert, Stahel, and Abbt 2001). The conceptual model underlying ACE holds that “just noticeable differences”—errors with an absolute size greater than \((w \times x_{i,t})\)—would not be tolerated or passed on. However, as a consequence of employing the normal distribution with mean 0 and variance \((w/2)^2\), approximately 4.55 percent of transmission events result in copying errors with absolute values that surpass \((w \times x_{i,t})\). Also problematic is the fact that drawing a value from the normal distribution makes it possible—although highly unlikely, given a small \(w\)—for \( x_{i,t+1} \) to take a negative value even though there is no such thing as a negative length, width, or thickness.

Hamilton and Buchanan (2009, Eq. 2) and Kempe, Lycett, and Mesoudi (2012) represent the same multiplicative random process with a slightly different time-series equation:

\[
x_{i,t+1} = x_{i,t} \times \left[ 1 + \mathcal{N}\left(0, w^2\right)\right],
\]

where \( \mathcal{N}(0, w^2) \) is a normal random variable with mean 0 and variance \(w^2\). Here, again, the Weber Fraction is used to parameterize a normally distributed error term. Because the variance of the normal distribution is greater in Eq. 2 than in Eq. 1, one can expect an even higher proportion—31.67 percent, to be exact—of transmission events to result in what would be visually perceptible deviations from the target value.

Although Eerkens and Lipo (2005), Hamilton and Buchanan (2009), and Kempe, Lycett, and Mesoudi (2012) study the same stochastic process with essentially the same time-series equation, they arrive at contradictory conclusions regarding the expected value—i.e., the mean—of the random variable \( X_{\infty,t} \) on the linear scale. According to Eerkens and Lipo (2005) and Kempe, Lycett, and Mesoudi (2012), the expected value of \( X_{\infty,t} \) remains constant with increasing \( t \) on the linear scale (see also Gardiner 2004, 103–4). For their part, Hamilton and Buchanan (2009) correctly show that the expected value of the natural logarithm of \( X_{\infty,t} \) decreases linearly with increasing \( t \) on the log scale at a rate of half the variance of the copying error term. These statements do not contradict each other because the former applies to Mean(\( X_{\infty,t} \)) on the linear scale while the latter applies to Mean(\( \ln X_{\infty,t} \)) on the log scale. But Hamilton and Buchanan (2009, 60) go on to argue that the expected value of \( X_{\infty,t} \) also decreases with increasing \( t \) on the linear scale at an exponentially decaying rate:

\[
E[X_{\infty,t}] = \bar{x}_{\infty,0} e^{-\alpha t}
\]

where \( \bar{x}_{\infty,0} \) represents the mean of \( X \) in an infinitely large population at \( t = 0 \) and \( \alpha \) is equal to the half the variance of the copying error term \((w^2/2)\) according to their model assumptions. These mutually exclusive predictions cannot both be correct. So, which is it: under vertical transmission, ought the expected value of \( X_{\infty,t} \) on the linear scale remain constant with increasing \( t \), or should it decrease at an exponentially decreasing rate with increasing \( t \)? It is crucial that we know the answer to this question before comparing empirical data to expectations generated by ACE. Think of how different one’s behavioral interpretation of the same empirical data set could be depending upon whether one uses the expectation that Mean(\( X_{\infty,t} \)) ought to remain constant under vertical transmission rather than the expectation that Mean(\( X_{\infty,t} \)) should decrease with increasing \( t \) under vertical transmission.

The key to understanding how the expected value of \( \ln X_{\infty,t} \) decreases at a constant rate on the log scale (as per Hamilton and Buchanan 2009) while the expected value of \( X_{\infty,t} \) remains constant on the linear scale with increasing \( t \) (as per Eerkens and Lipo 2005; Kempe, Lycett, and Mesoudi 2012) is to recall that the multiplicative stochastic process described in Eqs. 1 and 2 results in a log normally distributed random variable (\( X_{\infty,t} \)) with a constant mean but a variance that increases exponentially with increasing \( t \) (see Kempe, Lycett, and Mesoudi 2012; Limpert, Stahel, and Abbt 2001), assuming—as ACE does—that the value of the continuous trait can take any positive value. The fact that \( X_{\infty,t} \) is log normally distributed with a mean that remains constant but a variance that increases on the linear scale with increasing \( t \) explains why the expected value of \( \ln X_{\infty,t} \) decreases on the log scale with increasing \( t \). We can illustrate this important point by comparing two log-normal distributions—\( Y \) and \( Z \)—that share the same mean (2.73) but differ in variance (Var(\( Y \)) = 13.01 and Var(\( Z \)) = 37.87) on the linear scale. On account of its higher variance, \( Z \) is more positively skewed than \( Y \) on the linear scale. Despite the fact that both log-normal distributions, \( Y \) and \( Z \), display the same mean on the linear scale, Mean(\( \ln Y \)) is 0.50 while Mean(\( \ln Z \)) is only 0.01 on the log scale (Figure 1). In other words, despite the fact that \( Y \) and \( Z \) display exactly the same mean on the linear scale, Mean(\( \ln Z \)) is fifty times less than Mean(\( \ln Y \)) on the log scale because \( Z \) is more positively skewed on the linear scale.

Under vertical transmission, Hamilton and Buchanan (2009, 57) are correct that Mean(\( \ln X_{\infty,t} \)) decreases at a constant rate on the log scale with increasing \( t \). But this
occurs because $\text{Var}(X_{\infty,t})$ increases exponentially with increasing $t$ while $\text{Mean}(X_{\infty,t})$ remains unchanged on the linear scale (as shown by Kempe, Lycett, and Mesoudi 2012, 3), not because $\text{Mean}(X_{\infty,t})$ decreases on the linear scale at an exponentially decreasing rate with increasing $t$ (contra Hamilton and Buchanan 2009, 60). In sum, Eerkens and Lipo (2005) and Kempe, Lycett, and Mesoudi (2012) are correct: under the assumptions of ACE, $\text{Mean}(X_{\infty,t})$ remains constant on the linear scale. Eq. 3 should not be used to estimate the expected value of $X_{\infty,t}$ on the linear scale.

At the close of this brief review, it is important to reiterate that we should expect $\text{Mean}(X_{\infty,t})$ to remain constant and $\text{Var}(X_{\infty,t})$ to increase exponentially on the linear scale with increasing $t$ only in the very special case that (1) the population contains infinitely many vertical transmission chains (i.e., $N = \infty$) and (2) the continuous trait of interest can take any positive value. Both assumptions are violated by every known empirical case. It is imperative that we understand how violating these assumptions can affect population-level expectations of the mean and variance of a continuous trait before employing the expectations generated by ACE to infer the mechanisms of cultural transmission responsible for the values we observe empirically. In the remainder of the article, I use simulation experiments to systematically investigate the effects of violating each of the assumptions after introducing a new version of ACE that guards against the transmission of perceptible errors.

**METHODS: A NEW VERSION OF ACE**

All three of the studies reviewed above address the same multiplicative stochastic process with basically the same formal model. While the two more recent examples differ from Eerkens and Lipo (2005) in the size of the variance of the error term, all three allow for the transmission of noticeable differences resulting from errors with absolute values greater than $(w \times x_{i,t})$. Perceptibly erroneous attributes probably were fashioned by mistake and may have even entered the archaeological record, but according to the conceptual model that underlies ACE, any error with an absolute value greater than $(w \times x_{i,t})$ yields a noticeably aberrant value of $x_{i,t+1}$ that would not be passed on. The following paragraph introduces a new version of ACE that does not allow the transmission of so-called just-noticeable differences.

Consider a constant-sized population of $N$ individuals, each of whom displays a continuous cultural trait, $X$. Under vertical transmission, the continuous trait is transmitted from the “experienced generation” $t$ to the “naïve generation” $t + 1$ in each of $i = 1, 2, \ldots, N$ transmission chains, such that each novice learns from a different experienced practitioner. Although vertical cultural transmission is often associated with parent–offspring transmission, it applies more generally to conditions in which each member of the experienced generation transmits information to a different member of the naïve generation. These conditions can be met outside of parent–offspring transmission, such as when each skilled expert trains a different (and biologically unrelated) apprentice in the art of flintknapping or ceramics. Note that in the context of this cultural evolutionary model, the term “generation” does not refer to a real human generation of approximately twenty years but rather to the periodicity with which the trait is transmitted among individuals. How widely this periodicity varies among continuous traits is an interesting empirical question worthy of more research. As before, let $x_{i,t}$ represent the value of the continuous trait displayed by the individual in the $i$-th transmission chain after $t$ transmission events and let $x_{i,t+1}$ represent the value displayed by the individual in the $i$-th transmission chain after $t + 1$ transmission events. My version of ACE differs from previous versions in how it operationalizes visual perception error. Unlike previous studies, I use a uniform distribution to model visual perception error. My time-series equation for the transition of $x_{i,t}$ to $x_{i,t+1}$ in a single transmission chain is:

$$x_{i,t+1} = x_{i,t} \times [1 + U (-w, w)] \tag{4}$$

where $U(-w, w)$ represents a random variable drawn from a uniform distribution between $-w$ and $w$ with mean 0 and variance $w^2/3$. Because it is bound by $-w$ and $w$, the uniform distribution ensures that the absolute value of the error introduced per transmission event cannot exceed $(w \times x_{i,t})$. Note that Eqs. 1, 2, and 4 differ only in the variance and/or distribution of the proportional error term. While I am of the opinion that Eq. 4 hews more closely to the conceptual model of proportional copying error because it makes use of the uniform distribution (see also Eerkens and Bettinger 2001, 495), it is important to stress that my conclusions regarding the effects of finite population size and functional constraints do not depend upon this assumption. That is to say, the results would be qualitatively similar even if I had...
Before presenting the results of my experiments, it is useful to briefly illustrate how Eq. 4 works in a finite population. Figure 2 depicts the trajectories of $N = 200$ vertical transmission chains over $t = 1,000$ transmission events, where $w = 0.03$ and $x_{i0} = 10$ in each chain. Figure 2a, which replicates Figure 1b from Kempe, Lycett, and Mesoudi (2012), shows that $x_{it}$ follows a unique “path” through $t$ in each vertical transmission chain. The right-skewed distribution of $X_{200,1000}$ depicted in Figure 2b is to be expected for a multiplicative random process. Checking in with the very same simulation after $t = 100,000$ transmission events (Figure 2d, 2e) reveals that positive skewness increases with increasing $t$, also as expected. But contrary to the expectations that the mean of the continuous trait remains constant while the variance increases exponentially with increasing $t$, this simulation shows that Mean($X_{200,100000}$) is orders of magnitude lower than Mean($X_{200,0}$) and that Var($X_{200,t}$) does not increase exponentially with increasing $t$. The next section presents a much larger set of experimental results that help explain why this occurs in a finite population.

**RESULTS I: POPULATION SIZE AFFECTS THE MEAN AND VARIANCE OF A CONTINUOUS CULTURAL TRAIT**

**Vertical Cultural Transmission**

I begin by simulating the scenario in which continuous trait $X$ is vertically transmitted through $t = 200,000$ transmission events in each of $N$ transmission chains. To systematically investigate how finite population size alone affects the predictions of ACE, I simply vary $N$ ($N = 10, 50, or 100$) while satisfying the assumption that $X$ can take any positive value. Data are collected from 100 unique simulations of the stochastic model (Eq. 4) at each value of $N$ (Supplemental Code S1).

Recall that under vertical transmission the random variable $X_{\infty,200000}$ is distributed log-normally due to the multiplicative effects of proportional copying error. Because the random variable is right-skewed, one should expect a finite random sample of $X_{\infty,200000}$ to underrepresent values to the right of the expected value to a greater extent than it under-represents values to the left of the expected value. This systematic bias towards lower-than-average values yields sample estimates of the mean and variance that are lower than...
the actual mean and variance of $X_{\infty,200000}$. Put differently, for sufficiently high $t$ one should expect $\text{Mean}(X_{N,t})$ and $\text{Var}(X_{N,t})$ to increase with increasing $t$. The simulation results bear this out: for the case of $t = 200,000$ and $w = 0.03$, larger $N$ yield larger $\text{Mean}(X_{N,200000})$ and $\text{Var}(X_{N,200000})$ under vertical transmission (Table 1).

Because the downward bias is a consequence of sampling a log-normal distribution that becomes increasingly right-skewed (at an increasing rate) with increasing $t$, the effect of $N$ on $\text{Mean}(X_{N,t})$ and $\text{Var}(X_{N,t})$ strengthens as $t$ increases (e.g., Figure 2). Assuming that all vertical transmission chains start with the same value and that the continuous trait can take any positive value, for any combination of $N$ and $w$ there exists a threshold value of $t$ below which the sample estimate will not be biased towards lower values simply because the log-normal distribution of the random variable ($X_{\infty,t}$) is not (yet) sufficiently skewed for sampling bias to take effect. In general, the threshold value marking the lower boundary of what would constitute “sufficiently high $t$” increases with increasing $N$ and/or decreasing $w$. This suggests one would need to assess how large $t$ is on a trait-by-trait basis before deciding whether ACE’s predictions are appropriate.

Under vertical cultural transmission, members of different transmission chains do not interact with each other. Thus, one can reasonably treat the transmission chains in each of our simulated populations as though they represent $N$ independent observations of the random variable $X_{\infty,200000}$. As Hamilton and Buchanan (2009) point out, this means one can “linearize” the observed values by log-transforming them before calculating the mean and variance. I have applied this treatment to my data. Table 1 shows that decreasing $N$ does not bias the sample estimate of the expected value of $\ln X_{\infty,200000}$ ($27.70$), which under the conditions of my model (Eq. 4) and following Hamilton and Buchanan (2009, Eq. 6) is given by:

$$E [\ln X_{\infty,t}] = \text{Mean} (\ln X_{\infty,0}) - \left( \left( \frac{w^2}{3} / 2 \right) t \right) \tag{5}$$

Finite population size does not introduce sampling bias in this case because the log-transformed, or “linearized,” values are distributed normally on the log scale. Also consistent with Hamilton and Buchanan (2009, 57), my results show that under vertical transmission, $\text{Var}(\ln X_{N,t})$ increases with increasing $t$ on the log scale at a constant rate regardless of $N$ (Table 1).

### Unbiased, Prestige Biased, and Conformist Biased Transmission

The results of the preceding section show that log-transforming $X_{N,t}$ alleviates the sampling bias introduced by finite population size. Although this suggests that those who apply this treatment might safely ignore $N$ when studying cultural transmission in population-level data, it remains to be seen whether log-transformation ensures that one can confidently discern nonvertical cultural transmission from vertical cultural transmission in finite populations. To address this issue, here I extend my experimental design to collect data under each of four nonvertical mechanisms of cultural transmission: unbiased transmission, prestige biased transmission, median conformist transmission, and mean conformist transmission (Supplemental Code S2).

Under unbiased transmission, each member of the $t + 1$ generation “chooses” (i.e., samples) its “teacher” from the $t$-th generation at random and with replacement. By contrast, under prestige biased transmission all members of the $t + 1$ generation learn from the same individual, the one deemed the “most prestigious” member of the $t$-th generation. Following Eerkens and Lipo (2005), prestige is assigned arbitrarily—one member of the $t$-th generation is chosen at random to serve as the teacher for all members of the $t + 1$ generation. This process is repeated once per time step so that each randomly chosen prestigious individual “teaches” all of the members of just one generation. Under median conformist transmission, every member of the $t + 1$ generation attempts to copy the median of the values displayed by the $t$-th generation (with the “ties” that occur when $N$ is an even number broken randomly). Similarly, under mean conformist transmission, every member of the $t + 1$ generation attempts to copy the mean of the values displayed by the $t$-th generation. Table 2 summarizes the results of 100 simulations run for each value of $N$ (10, 50, or 100) under each of the four nonvertical mechanisms of cultural transmission. Just as observed above in the case of vertical transmission, in finite populations $\text{Mean}(X_{N,t})$ does not remain constant nor does $\text{Var}(X_{N,t})$ increase monotonically on the linear scale with increasing $t$ under unbiased, prestige biased, or conformist transmission (Table 2).

Applying Hamilton and Buchanan’s (2009) treatment to the data collected under nonvertical transmission raises a number of interesting points regarding prestige biased and unbiased transmission. First, the results I obtain under
prestige biased transmission are qualitatively similar to those Hamilton and Buchanan (2009) obtained from their BACE model: Mean(ln $X_{t,N}$) decreases through time while Var(ln $X_{t,N}$) asymptotes at an equilibrium value. Second, the log-transformed data suggest that under unbiased cultural transmission Mean(ln $X_{t,N}$) decreases more slowly when $N$ is larger due to a larger effective population size ($N_e$) (Premo 2016). Third, unbiased and prestige biased transmission yield similar Mean(ln $X_{t,200000}$) values for the population sizes tested here. What is even more concerning, the Mean(ln $X_{t,200000}$) values observed in finite populations under unbiased and prestige biased transmission overlap considerably with those observed under vertical transmission (compare Tables 1 and 2) as well as with the expected value of ln $X_{\infty,200000}$ (-27.70). This degree of equifinality strongly suggests it would be difficult—perhaps even unwise—to use Mean(ln $X_{t,N}$) alone to discern among vertical, unbiased, and prestige biased transmission in finite populations for cases in which one suspects $t$ is sufficiently high for a given $N$.

Table 2 shows that, for sufficiently high $t$, $N$ has a much weaker effect on the rate at which Mean($X_{t,N}$) and Mean(ln $X_{t,N}$) decrease with increasing $t$ under either form of conformity. Conformist transmission allows Mean($X_{t,N}$) and Mean(ln $X_{t,N}$) to “wander” more freely through time than is possible under unbiased, prestige biased, or vertical transmission in a finite population. On the other hand, compared to unbiased or vertical transmission, conformist and prestige biased transmission severely limit the population-level variance at any single point in time. Because under conformist transmission every member of the $t + 1$ generation attempts to copy a single value (i.e., the mean or the median of the $t$-th generation), Var(ln $X_{t,200000}$) closely approximates $w^2/3$, which not coincidentally is the variance of the uniform distribution used to represent visual perception error (Eq. 4). Table 2 shows that prestige biased transmission has the same effect on Var(ln $X_{t,200000}$) for a similar reason: every member of the $t + 1$ generation attempts to copy the value displayed by the single most prestigious individual of the $t$-th generation. A practical implication of this finding is that Var(ln $X_{t,N}$) alone cannot distinguish prestige from conformist biased transmission regardless of $N$ or $t$.

**RESULTS II: FUNCTIONAL CONSTRAINTS STABILIZE THE MEAN AND VARIANCE OF A CONTINUOUS CULTURAL TRAIT**

Eerkens and Lipo (2005) offer “length of an arrowhead” as an example of the kind of real-world continuous trait to which ACE applies. Although this suggestion is helpful conceptually, strictly speaking, the kind of continuous trait ACE models is not analogous to the length of an arrowhead simply because ACE allows the continuous trait to take any positive value. Without a “ceiling,” $x_{t,N}$ can increase without bound with increasing $t$, potentially resulting in astronomically large values given enough transmission events. If $X$ can represent the length of a projectile point, then this assumption simply does not fly. Just as problematically, although ACE also allows $x_{t,N}$ to take infinitesimally small values, one cannot use—let alone manufacture by hand—a projectile point that is only 0.001 cm long.

ACE’s assumption that a continuous trait can take any positive value is untenable on principle alone. The metric attributes of all handmade implements are constrained at both the lower and upper bounds by the dimensions of the human body, the tool’s function, and the physical properties of the raw material(s) from which the tool is made. A 200-centimeter-long projectile point is of little use to a hunter pursuing deer, and thus would not be transmitted to a novice. The same is true of an equally useless 0.001-centimeter-long projectile point. Despite the term’s connotation, functional constraints also apply to purely symbolic traits. Surely there are “functional,” or perhaps it is better to say “physical,” constraints on the length of a strictly symbolic pendant worn around one’s neck. A 200-centimeter-long stone pendant would be far too awkward to wear on a necklace, and a 0.001-centimeter-long pendant would be difficult to see with the naked eye. Although the degree to
which functional constraints vary among continuous traits is an interesting and important empirical question, the inescapable fact that all metric attributes of handmade implements are limited to some extent raises the question: are ACE’s population-level expectations useful despite the fact that they depend upon the assumption that a continuous trait can potentially take any positive value, no matter how large or small?

Consulting the cultural evolution literature regarding this issue, one quickly encounters the notion that functional constraints ought to have little effect on how cultural variation accumulates within and between groups in a structured population. Richerson and Boyd (2005, 142; emphasis added) state “models predict that traditions among small, semi-isolated groups will rapidly diverge, so that even if functional constraints are strong, variation will increase between groups through time.” While this statement is true for a selectively neutral discrete trait that can be expressed as one of infinitely many different variants (according to the “infinite variants” model of copying error) by members of isolated groups, I am skeptical that it holds more generally (nor do I think Richerson and Boyd imply that it ought to), such that it also applies to a continuous trait that is functionally constrained similarly across groups. Yet, in the context of explaining how proportional copying error ought to affect continuous attributes of Acheulean handaxes, Corbey et al. (2016, 9; emphasis added) echo Richerson and Boyd’s passage in stating “over the course of almost 1.5 million years, 3%–5% copying errors would have resulted in a substantial amount of spatiotemporal variation, regardless of how closely linked handaxes were to a particular function.” I could find no support in the cultural evolution literature for the assertion that the variation-enhancing effects of proportional copying error will necessarily outweigh the variation-limiting effects of functional constraints. In fact, to the best of my knowledge, no one has studied a version of ACE that imposes functional constraints on a continuous trait. Mesoudi and Lycett’s (2009) investigation of the effects of frequency-dependent trait trimming—a form of cultural transmission biased against very rare or very common discrete variants—perhaps comes the closest. But because they focus on the relative frequencies of different variants of a discrete trait rather than the mean and variance of a continuous trait, their model does not include the effects of proportional copying error that are central to ACE, and, thus, it is not germane to the issue addressed here.

Here, I further modify the model introduced above (Eq. 4) to investigate how functional constraints affect the population-level mean, variance, and coefficient of variation (CV; calculated as the sample standard deviation divided by the sample mean) of a continuous trait passed via vertical cultural transmission. FACE (Functional constraints + ACE) differs from the version of ACE employed in the previous section in one important way: during the transition from \( x_{1:t} \) to \( x_{1:t} + 1 \) FACE enforces the rule that \( x_{1:t} + 1 \) must fall within an arbitrarily defined range of functionally relevant values. More specifically, whenever Eq. 4 yields an \( x_{1:t} + 1 \) less than 5 or greater than 10, FACE forces a “retry” until \( x_{1:t} + 1 \) falls within this arbitrarily defined range of acceptable values (Supplemental Code S3).

Let us begin by considering the population-level mean, variance, and CV one would expect to see in a continuous attribute (e.g., length) of a handmade implement (e.g., stone projectile point) under the “null” condition in which we assume that everyone in the population independently discovers a functional value of the attribute via individual learning rather than through cultural transmission. Notice that in this scenario, the attribute is not a cultural trait, per se, because it is not transmitted; hence, my purposeful use of the term “continuous attribute” rather than “cultural trait.” Because the attribute is not transmitted from an experienced practitioner to a novice but rather independently discovered by each novice through individual learning, there is no copying error.

In the absence of cultural transmission—and, more to the point, in the absence of the proportional copying error it introduces—the mean, variance, and CV of the individually learned continuous attribute are no longer functions of \( w, N, \) or \( t \), and, thus, there is no reason to expect that the mean, variance, or CV of the attribute ought to change with increasing \( t \). Under the “null” (asocial) condition described above, the expected mean, variance, and CV of the continuous attribute in a population of \( N \) toolmakers are simply given by the mean, variance, and CV of a uniform distribution bound by the lower and upper limits to the trait’s functionally relevant values. The use of a uniform distribution allows each novice to arrive at the solution that works best for that individual—given the size or strength of one’s hands, for example—within the wider range of functionally relevant values. Assuming for the purposes of this exercise that the functionally relevant values of the attribute of interest are bound by 5 and 10, one would expect the distribution of projectile point lengths displayed by a population of individual learners to show a mean of 7.5, a variance of 2.08, and a CV of 0.19 regardless of \( N \) or \( t \). On the log scale, one would expect a mean of 2.01, a variance of 0.04, and a CV of 0.10 regardless of \( N \) or \( t \).

With these “null” expectations in hand, we can turn to FACE to simulate the scenario in which we assume that projectile point length is passed from experienced to naive individuals via vertical cultural transmission. Like ACE, FACE includes proportional copying error. But unlike ACE, FACE imposes functional constraints that ultimately limit the range of possible trait values to between 5 and 10. Table 3 provides data collected from 100 populations consisting of \( N = 10, 50, \) or 100 vertical transmission chains simulated over 1,000 transmission events in the presence of functional constraints (Supplemental Code S3). The results show that the mean, variance, and CV of \( X_{5:1000} \) and \( \ln X_{5:1000} \) are strongly affected by functional constraints even in the presence of proportional copying error. Second, and just as interestingly, for all population sizes tested here, FACE
yields mean, variance, and CV values on both the linear and log scale that are only slightly different from what one would expect of a continuous attribute in the absence of cultural transmission altogether (Table 3).

The take-home message of Table 3 is that the selective pressure supplied by functional constraints severely weakens the cumulative effect of proportional copying error on the population-level mean, variance, and CV of a continuous trait regardless of population size. In fact, the effect of proportional copying error is so weak relative to the effect of functional constraints that it will likely be very difficult, if not impossible, to distinguish cultural transmission from individual learning given the population-level mean, variance, or CV of any continuous trait constrained by a rather narrow range of functionally relevant values. At this point it is worth asking oneself which continuous attributes of handmade technologies are not marked by a prohibitively narrow range of functional values? Future applications should focus on such traits, if there are any, rather than merely settling for the traits made available to us by the vagaries of time and taphonomy. It is important to acknowledge that there is no work-around for the variation-limiting effects of functional constraints. My results show that log-transformation does not, and indeed cannot, mitigate the effect of functional constraints, even under vertical transmission (Table 3).

The implication of this finding is clear: employing expectations that ignore the effects of a trait’s functional constraints increases the likelihood of incorrectly interpreting a relatively stable mean or an ostensibly “lower-than-expected” level of variation as evidence of biased cultural transmission when in fact the trait of interest was transmitted vertically—or perhaps not even transmitted at all, but rather discovered through individual learning. This will not come as news to Kempe, Lycett, and Mesoudi (2012, 6), who state that “functionally-related cultural selection,” driven by the basic need for a tool to fit in its user’s hand, for example, belongs on the list of possible explanations for the ostensibly lower-than-expected CV in handaxe size. Given its ubiquity, it would seem that functionally related stabilizing selection generally provides a more parsimonious explanation than conformist biased transmission for “lower-than-expected” variation in a continuous trait.

It is also worth pointing out that by stabilizing the mean and variance of a continuous trait, functional constraints also limit the amount of continuous cultural variation that can accumulate between semi-isolated, or even completely isolated, groups in a structured population. If the continuous cultural trait of interest is constrained similarly wherever and whenever it is displayed, then, contrary to the assertion by Corbey et al. (2016) quoted above, one should expect low between-group variation despite the effects of proportional copying error and even in the absence of intergroup transmission.

What is more, the stabilizing effect of functional constraints opens the door to potential explanations of stasis and “lower-than-expected” continuous variation that do not require social learning at all. The “null” condition represented in the thought experiment above shows that when the value of a continuous attribute is physically constrained by its function or by the materials or techniques used to make it, individual learning alone yields a stable mean and relatively low variance and CV. Holding all else constant, a continuous trait marked by a narrow range of functionally relevant values will display a relatively stable mean as well as low variance and CV regardless of whether it is learned socially or individually. Thus, a continuous trait that displays a stable mean, low variance, or low between-group variation does not necessarily indicate the presence of wide-scale conformist biased transmission. In fact, it may be indicative of a technological attribute that was learned individually rather than socially.

**DISCUSSION**

Previous applications of ACE provide expectations for the population-level mean and variance of a vertically transmitted continuous cultural trait. These expectations depend upon two important assumptions: (1) the population has infinitely many vertical transmission chains; and (2) the continuous trait of interest can take any positive value. This article employs simulation experiments to investigate the consequences of violating either assumption. The results teach two fundamental lessons that hold important implications for how evolutionary anthropologists interpret empirical departures from the predictions of ACE.

First, population size matters, and it matters more and more with increasing t. Holding w and N constant, the mean...
and variance of $X_{N,t}$ decrease on the linear scale with increasing $t$ due to the effects of proportional copying error and sampling bias. This result is explained by the fact that under vertical cultural transmission any finite population serves as a small, random sample of the log-normally distributed random variable $X_{\infty,t}$, which is characterized by a stationary mean and an exponentially increasing variance (on the linear scale). The fact that $X_{\infty,t}$ is log-normally distributed on the linear scale explains why a finite sample of vertical transmission chains does not bias estimates of the mean and variance of $\ln X_{\infty,t}$ on the log scale.

In the case of vertical transmission, log-transforming $X_{N,t}$ before calculating the observed mean and variance effectively mitigates the sampling bias introduced by finite $N$. This treatment works well for simulated data that we know were collected under vertical cultural transmission. But empirically we do not know a priori whether a trait was passed via vertical transmission—in fact, that is the question ACE was designed to help us address. Unfortunately, under conditions marked by sufficiently high $t$, unbiased and prestige biased transmission yield Mean($\ln X_{N,t}$) that are very similar to the expected values of both $\ln X_{N,t}$ and $\ln X_{\infty,t}$ under vertical transmission. This means that even after log-transformation it remains difficult to discern among vertical, unbiased, and prestige biased transmission from Mean($\ln X_{N,t}$) alone. Here, an additional line of evidence—Var($\ln X_{N,t}$)—provides some inferential leverage. Although Var($\ln X_{N,t}$) does not distinguish prestige bias from conformity, one ought to expect much lower Var($\ln X_{N,t}$) under unbiased, prestige biased, and conformist cultural transmission than under vertical transmission. But even this holds true only under the untenable assumption that the continuous trait of interest can take any positive value.

The second lesson is that any continuous trait constrained to a relatively narrow range of functionally relevant values provides a particularly poor choice for studying cultural transmission by way of comparing population-level observations against expectations derived from the cumulative effects of proportional copying error. The fact that every metric attribute of a handmade implement, regardless of technology or period, violates the assumption that it can become infinitely large raises the specter that previous incarnations of ACE systematically overestimate the population-level variation one ought to expect under vertical cultural transmission. There are good commonsense reasons to assume that every continuous attribute of material culture has nonzero lower and upper bounds that mark the physical constraints imposed by tool manufacture and use. And remember that although the term “functional” conjures images of implements used to procure calories, provide shelter, raise offspring, secure mates, or defend resources, functional constraints also limit continuous attributes of purely symbolic items (Dunnell 1978).

FACE illustrates how functional constraints increase the likelihood of convergence in continuous trait values (e.g., O’Brien, Buchanan, and Eren 2018), regardless of population size and regardless of how, or even whether, the trait was culturally transmitted between individuals. The stabilizing effect of functional constraints will result in low variation both within and between isolated groups regardless of whether the continuous attribute of interest is transmitted culturally or learned individually, so long as the trait is constrained similarly across groups. Explicitly incorporating functional constraints into future models should help us avoid uncritically accepting similarities between what may in fact be analogous—as opposed to homologous—trait values in the Paleolithic archaeological record as evidence of a shared cultural “tradition” spanning multiple continents and persisting over hundreds of thousands of years and multiple fossil hominin species. Likewise, we would be wise to account for the effects of functional constraints when evaluating claims that individual learning, rather than conformist or prestige biased cultural transmission, may be sufficient to explain what appear to be strikingly slow rates of change and low levels of spatiotemporal metric variation in Early Stone Age tools (see Ambrose 2001; Richerson and Boyd 2005; Tennie et al. 2017).

**CONCLUSION**

Hamilton and Buchanan (2009) are correct that the expected value of the natural logarithm of $X_{\infty,t}$ decreases linearly on the log scale with increasing $t$. Eerkens and Lipo (2005) and Kempe, Lyckett, and Mesoudi (2012) are correct that the expected value of $X_{\infty,t}$ does not change on the linear scale with increasing $t$. Although these statements seem contradictory, they can both be true because, under the assumptions of ACE, the variance of $X_{\infty,t}$ increases exponentially on the linear scale (and linearly on the log scale) while its mean remains constant on the linear scale (but decreases linearly on the log scale) with increasing $t$. However, the fact that these expectations hold only for special “idealized” conditions that are violated by all empirical cases should cause pause for concern. My results suggest that previous applications of ACE likely overestimate the expected value and variance one ought to see in a finite population under vertical transmission, thereby increasing the likelihood of unnecessarily invoking biased cultural transmission to explain “lower-than-expected” observed values on the linear scale. Hamilton and Buchanan (2009) present a work-around for the case of any trait known to have been transmitted via vertical transmission: log-transforming the observed values before calculating the mean and variance alleviates the bias inherent to randomly sampling what is ultimately a log-normally distributed random variable. But this option comes with some strings attached: (1) when $t$ is sufficiently high, unbiased and prestige biased cultural transmission both yield Mean($\ln X_{N,t}$) that are very similar to the expected value of $\ln X_{\infty,t}$ under vertical transmission; and (2) Var($\ln X_{N,t}$) does not distinguish conformist from prestige biased transmission regardless of $N$ or $t$.

The sampling bias introduced by finite population size also compromises studies that compare observed values to
ACE’s expectations on the linear scale (Eerkens and Lipo 2005; Kempe, Lycett, and Mesoudi 2012). My findings illustrate why one must be mindful of $N$ and $t$ when comparing $\text{Mean}(X_{N,t})$, $\text{Var}(X_{N,t})$, or $\text{CV}(X_{N,t})$ to the expectations of $\text{Mean}(X_{\infty,t})$, $\text{Var}(X_{\infty,t})$, or $\text{CV}(X_{\infty,t})$ provided by ACE. Under ACE’s assumption that a continuous trait can take any positive value, for sufficiently high $r$ a very large $N$ is required to incorporate enough of the relatively large but rare values needed to “counterbalance” the more numerous but smaller values in the log-normally distributed random variable $X_{\infty,t}$. Assuming that all transmission chains in the constant-sized population are the same “length” (i.e., assuming that $t$ does not vary among transmission chains) and that all chains were seeded with the same value at $t = 0$, just how large $N$ must be to avoid sampling bias depends upon the values of $r$ and $w$. Perhaps ACE could be of some utility in empirical cases where independent lines of evidence suggest $r$ is relatively low and $N$ is relatively high, assuming of course that the trait of interest also has a very large range of functionally relevant values. However, in cases where the number of transmission events per chain remains unknown, it may be best practice to simply avoid comparing the observed population-level mean and variance to the expected mean and variance of $X_{\infty,t}$ on the linear scale.

Under the assumption that the continuous trait of interest can take any positive value, ACE predicts the vast majority of values in the log-normally distributed random variable will be lower than is physically possible for a metric attribute of any handmade implement. But assuming that a continuous trait has no functional constraints causes problems beyond allowing for projectile points that are only 0.0001 cm long. Functional constraints limit the amount of variation that would otherwise accumulate with increasing $t$ due to the multiplicative effects of proportional copying error. Thus, in the presence of functional constraints, one should expect the mean and variance of a continuous trait to remain constant with increasing $t$ (barring a shift in the range of functionally relevant values) regardless of the mechanism of cultural transmission and despite the effects of proportional copying error. In the end, ignoring the effects of functional constraints further increases the likelihood of falsely invoking biased cultural transmission to explain a stable mean or low variance in a finite population. What is more, FACE shows that in the presence of functional constraints one ought to expect the mean, variance, and CV of a metric attribute to be unaffected by whether it is obtained via cultural transmission or individual learning. As a practical matter, these findings strongly suggest to me that individual learning rather than high-fidelity cultural transmission provides a more parsimonious explanation for the observation that quantitative attributes of Early Stone Age implements show evidence of what appear to be very slow and nondirectional diachronic change and relatively little regional differentiation—although it must be noted that these archaeological signals themselves might be undesirable artifacts of “significant time-averaging” (Kowalewski 1996) rather than trustworthy indicators of the behavior displayed by any Paleolithic population, let alone all of them (e.g., Miller-Atkins and Premo 2018).

All of this leads to the conclusion that, in lieu of independent evidence of $r$, $N$, and the continuous trait’s functional constraints, ACE’s predictions provide less inferential power than previously thought, even after log-transforming population-level observations. Despite this, ACE still might provide a foundation upon which we can build a better model. To that end, we should invest in developing generative models, like FACE, that explicitly consider the functional constraints of continuous traits in addition to the psychophysical limits of the individuals who transmit them. There is reason to be optimistic on this front. It should be possible to incorporate reasonable estimates of functional constraints in light of basic information regarding how an implement was made and for what tasks it was likely to have been used. At the very least, we can probably agree that a handaxe cannot be smaller than required to hold between one’s fingers nor larger than could be carried by an individual. Or perhaps we could go one step further and use the smallest known handaxe and largest known handaxe as guides for our lower and upper bounds on handaxe size. Although such priors are admittedly coarse-grained, they should be included in formal models that generate expectations of the mean and variance in handaxe size in finite populations. Surely even those rough estimates are preferable to an assumption that allows handaxes to become larger than Montana or smaller than a microbe.

The next generation of accumulated copying error models will benefit from detailed ethnographic, ethnoarchaeological, and experimental archaeological investigations into the functional constraints that mark continuous attributes of material culture. And yet, I suspect that incorporating experimentally informed estimates of functional constraints in future models is likely to reveal even more fully that the task of distinguishing purely individual learning from high-fidelity cultural transmission—let alone discerning among vertical, unbiased, prestige biased, and conformist cultural transmission—from population-level data (not to mention assemblage-level data) is more fraught than evolutionary anthropologists and archaeologists currently appreciate. The results of my study suggest the task will be most difficult for Paleolithic archaeologists, who due to no fault of their own are relegated to continuous attributes of stone tools (most of which are marked by a relatively narrow range of functionally relevant values) that are presumed to have been transmitted over tens to hundreds of thousands of years (sufficiently high $t$) within relatively small populations (low $N$).

L. S. Premo
Department of Anthropology, Washington State University, PO Box 4910, Pullman, WA, 99164, USA and Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig, 04103, Germany, luke.premo@wsu.edu
I am lucky to have friends like Dr. Anne Kandler and Prof. Dick Gumulkiwitz, who were kind enough to spend time discussing this project with me over the last three years. I thank Anne for teaching me about Jensen’s Inequality and for commenting on a very early draft of this paper. I thank Prof. Jean-Jacques Hublin and Prof. Richard McElreath for supporting my sabbatical stay at the Max Planck Institute for Evolutionary Anthropology; I could not have completed this project anywhere else. I thank Dr. Shannon McPherron, Dr. Fabrizio Mafessoni, and five anonymous reviewers for their constructive criticism and helpful comments. Their feedback improved this paper considerably.

1. To be precise, Hamilton and Buchanan (2009) also present another version of ACE that differs further in that it includes a form of “structural error” that is independent of the target value. But this additive (rather than multiplicative) form of error does not appear to change the dynamics of ACE.

**Data Availability Statement**

The fully commented R source code for generating the data analyzed in this paper is provided in the electronic supplemental materials. The source files are also freely available upon request from the author.

**REFERENCES CITED**

Ambrose, Stanley H. 2001. “Paleolithic Technology and Human Evolution.” *Science* 291:1748–53.

Corbey, Raymond, Adam Jagich, Krist Vaesen, and Mark Collard. 2016. “The Acheulean Handaxe: More Like a Bird’s Nest Than a Beatles’ Tune?” *Evolutionary Anthropology* 25:6–19.

Coren, Stanley, Lawrence M. Ward, and James T. Enns. 1994. *Sensation and Perception*. San Diego: Harcourt Brace.

Dunnell, Robert C. 1978. “Style and Function: A Fundamental Dichotomy.” *American Antiquity* 43 (2): 192–202.

Eerkens, Jelmer W. 2000. “Practice Makes Within 5% of Perfect: Visual Perception, Motor Skills, and Memory in Artifact Variations.” *Current Anthropology* 41 (4): 663–68.

Eerkens, Jelmer W., and Carl P. Lipo. 2005. “Cultural Transmission, Copying Errors, and the Generation of Variation in Material Culture and the Archaeological Record.” *Journal of Anthropological Archaeology* 24:316–34.

Eerkens, Jelmer W., and Robert L. Bettinger. 2001. “Techniques for Assessing Standardization in Artifact Assemblages: Can We Scale Material Variability?” *American Antiquity* 66 (3): 493–504.

Gardiner, Crispin W. 2004. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Third edition. New York: Springer.

Gilinsky, Albert S. 1951. “Perceived Size and Distance in Visual Space.” *Psychological Review* 58 (6): 460–82.

Hamilton, Marcus J., and Briggs 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.

Kempe, Marius, Stephen Lycett, and Alex Mesoudi. 2012. “An Experimental Test of the Accumulated Copying Error Model of Cultural Mutation for Acheulean Handaxe Size.” *PLoS One* 7 (11): e48333 https://doi.org/10.1371/journal.pone.0048333.

Kowalewski, Michal. 1996. “Time-Averaging, Overcompleteness, and the Geological Record.” *The Journal of Geology* 104:317–26.

Limper, Eckhard, Werner A. Stahel, and Markus Abb. 2001. “Log-Normal Distributions across the Sciences: Keys and Clues.” *Biological Science* 51 (5): 341–52.

Mesoudi, Alex. 2011. *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*. Chicago: University of Chicago Press.

Mesoudi, Alex, and Stephen J. Lycett. 2009. “Random Copying, Frequency-Dependent Copying, and Culture Change.” *Evolution and Human Behavior* 30:41–48.

Miller-Atkins, Galen, and L. S. Premo. 2018. “Time-Averaging and the Spatial Scale of Regional Cultural Differentiation in Archaeological Assemblages.” *Science & Technology of Archaeological Research* 4 (1): 12–27. https://doi.org/10.1080/20548923.2018.1504490.

Norwich, Kenneth H. 1987. “On the Theory of Weber Fractions.” *Perception & Psychophysics* 42 (3): 286–98.

O’Brien, Michael J., Briggs Buchanan, and Metin I. Eren, eds. 2018. *Convergent Evolution in Stone-Tool Technology*. Cambridge: MIT Press.

Ono, Hiroshi. 1967. “Difference Threshold for Stimulus Length under Simultaneous and Non-Simultaneous Viewing Conditions.” *Perception & Psychophysics* 2:201–07.

Premo, L. S. 2016. “Effective Population Size and the Effect of Demography on Cultural Diversity and Technological Complexity.” *American Antiquity* 81 (4): 605–22.

Richerson, Peter J., and Robert Boyd. 2005. *Not By Genes Alone: How Culture Transformed Human Evolution*. Chicago: University of Chicago Press.

Tennie, Claudio, L. S. Premo, David R. Braun, and Shannon P. McPherron. 2017. “Early Stone Tools and Cultural Transmission: Resetting the Null Hypothesis.” *Current Anthropology* 58 (5): 652–72.

**Supporting Information**

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supplemental Code S1

Supplemental Code S2

Supplemental Code S3