Time-averaging and the spatial scale of regional cultural differentiation in archaeological assemblages

Galen Miller-Atkins\textsuperscript{a} and L. S. Premo\textsuperscript{a, b}

\textsuperscript{a}Department of Anthropology, Washington State University, Pullman, WA, USA; \textsuperscript{b}Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

ABSTRACT

The degree to which societies differ in dress, diet, laws, and language appears to be such an integral part of today’s human experience that some researchers think of it as a hallmark of so-called “modern human behavior.” Yet it remains unclear to what extent the current pattern of relatively low within-region cultural variation paired with relatively high between-region cultural variation can be assessed in time-averaged Paleolithic assemblages. Here, we use a spatially explicit agent-based model to begin to examine how time-averaging can affect the spatial scale of similarity among culturally transmitted variants in archaeological assemblages. Our results show that time-averaging, alone, can increase the scale of local spatial association among the relative frequency of the most prevalent cultural variant in an archaeological landscape. Our findings have important implications for archaeological interpretations of the spatial scale of regional cultural differentiation (or lack thereof) in the Paleolithic record and beyond.

1. Introduction

Human cultural variation is not randomly distributed over space. People in the same geographic region tend to share a greater number of norms, traditions, and languages with each other than they do with people chosen randomly from distant regions. Today’s spatially balkanized cultural variation is the result of a confluence of cultural evolutionary forces (including vertical and conformist biased cultural transmission), political history, and isolation by geographic distance. These processes have presumably both increased variation among geographically defined human populations and dampened variation within them. Although we are accustomed to a world in which a larger proportion of the total cultural variation is explained by cultural variation between societies than by the variation found within them, we do not know the antiquity of this level of regional cultural differentiation.

The purported appearance of culturally distinct, regional societies – as evidenced by the balance of variation within and between stone tool assemblages separated by hundreds or even thousands of kilometers – has been used as an archaeological signal of the presence of so-called “modern human behavior,” defined as “behavior that is mediated by socially constructed patterns of symbolic thinking, actions, and communication that allow for material and information exchange and cultural continuity between and across generations and contemporaneous communities” (Henshilwood and Marean 2003, 635). Assuming that the form of a discarded stone tool (or the technology used to make it) directly reflects culturally transmitted information, the spatial extent of similar-looking (or similarly-made) stone tools is thought to offer a proxy for the spatial extent of the society that produced, used, and discarded them.

Regional cultural differentiation is generally identified by both an increase in the number of culturally-distinct populations – societies, to use a more familiar term – and a decrease in the spatial scale of each society. Thus, archaeologists have attempted to address the process of regional cultural differentiation by looking for signs of the fragmentation of a large culturally-homogeneous entity into smaller internally-homogeneous but externally-heterogeneous units. Regionally distinctive cultures may date to as early as the Middle Stone Age (MSA) (Clark 1988; Klein 2008; McBrearty and Brooks 2000; Marean 2015; Wurz 2013). Although Clark (1988) is careful to acknowledge serious limitations imposed by a small sample size and an imprecise chronology, he tentatively suggests that differences in the “Heavy Duty” components of MSA assemblages recovered from sites in three different “zones” of East Africa – the Lake Victoria Basin, Eastern Sudan, and the humid coastal zone – may signify an early
development of regional cultures or identities. Wurz (2013, S312) suggests that a unique ~42,000-year-old assemblage from Border Cave “may be an indication that regionally distinct technological trajectories existed” in South Africa by the Late Stone Age (LSA). Archaeological research into the process of regional cultural differentiation is not restricted to the Paleolithic. Porcic and Nesic (2014) and Porcic (2015) find that the spatial extent of Neolithic cultures in the Balkans, primarily defined by variation in the relative frequencies of pottery decorations, may be explained by geography and unbiased cultural transmission.

Here it is important to note that although regional cultural differentiation is a population-level characteristic, an emergent property of how ideas are passed between social learners in a spatially explicit setting, archaeologists unfortunately do not have direct access to past populations. Instead, archaeologists study time-transgressive assemblages of artifacts. Whereas the term “population” refers to the set of individuals alive at a single point in time (one can talk about the current population of New Guinea, for instance), Paleolithic archaeological assemblages are very often (although not always) heavily time-averaged (Stern 1993, 1994). Time-averaging is the phenomenon whereby remains deposited at different times appear to the researcher to be pene-contemporaneous (Kowalewski 1996; Staff et al. 1986; Stern 1993, 1994; Terry 2008). An assemblage of stone tools recovered from a time-averaged layer may appear to have been deposited at roughly the same time – perhaps even by the same group of people – when in fact they were deposited at different times by different people during the formation of the stratigraphic layer over thousands of years (Binford 1981; Stern 1993, 1994). While it is clear that stone tools in such an assemblage do not necessarily represent the remains of any single past population (i.e. a group of people alive at the same time), it is often not very clear what they do represent.

Time-averaging can have interesting and non-intuitive effects on assemblage data. According to Kowalewski (1996), while time-averaging can help reduce the “noise” associated with short-term deviations from a long-term behavioral signal, it can also generate a false signal. This important point requires some unpacking. When the temporal scale of the behavior of interest is shorter than the duration of a time-averaged deposit, time-averaging reduces the temporal resolution of the data such that the assemblage reflects an “average” of the behaviors exhibited over a long period. Some archaeologists have argued that this may in fact be a good thing; that by buffering against the “noise” associated with daily, weekly, or monthly idiosyncrasies that are less likely to be important to evolutionary explanations of human behavior over the long durée, time-averaging might provide one with a clear view of the modal, or most consistently displayed, hominin behavior for a given period (Bailey 1983, 2007; Wandsnider 1996).

Following this line of thought archaeologists sometimes characterize time-averaging as a filter capable of distilling long-term behavioral signals to their essence without distorting them, but this is not necessarily the case. Time-averaging can degrade rather than purify the behavioral signal. The degree to which time-averaging degrades – i.e. averages and, thus, distorts – the signal of past behavior in an assemblage, which itself is a composite of synchronic and diachronic behavioral variation, depends on the difference between the time-scale of the behavior of interest and the duration of assemblage formation. As Kowalewski puts it:

Because the consequences of time-averaging depend entirely on the time-scale of a process (i.e. the time-averaging threshold), the identification of this timescale is fundamental. If the time-averaging threshold of a process is not well understood, then it is difficult to decide whether time-averaging eliminated the noise or averaged the signal. [Kowalewski 1996, 325]

Comparing the degree to which many Paleolithic assemblages are time-averaged (e.g. Stern 1994) to the relatively short time scales associated with the kinds of “processes” that interest archaeologists, like teaching someone how to make a useful implement out of stone, wood, or bone; the use-life of a tool; or even the fate of a group of hominin foragers, strongly suggests that Paleolithic assemblages regularly average the archaeological signals associated with hominin behaviors that are marked by time-averaging thresholds lower than the span of a single generation. A culture, or society, is a dynamic, time-transgressive, emergent entity that nevertheless persists over many generations before shifting (through cultural evolutionary processes) into something that looks different given the benefit of hindsight. While the time-averaging threshold associated with the rise and fall of a society is certainly higher than those associated with the transmission of knowledge from an experienced individual to her pupil or the use-life of a stone tool, the threshold is still likely to be lower than the thousands of years represented in many time-averaged Paleolithic assemblages (Stern 1993, 1994).

Finally, and perhaps even more concerning, time-averaging can also introduce patterns in assemblage data that are not representative of any past population. The “false signals” of which Kowalewski warns are emergent properties of assemblages that are not directly representative of either short- or long-term behaviors displayed by any of the populations that contributed material to the time-averaged record (see Premo (2014) for an illustration of this effect on cultural diversity).

In this study, we employ a spatially explicit agent-based model as a heuristic tool to assess how time-averaging can impact one’s ability to infer the true scale of
regional cultural differentiation in past populations from the archaeological assemblages they left behind. We find there are many conditions in which time-averaging alone can increase the scale of local spatial association in the relative frequency of the most prevalent cultural variant found in the archaeological landscape. In other words, we find that heavily time-averaged assemblages may not necessarily reflect the true spatial scale of regional differentiation exhibited by any of the past populations that deposited artifacts to the record. One should expect the signal observed in the assemblage to depart from the true scale when the minimum archaeological-stratigraphic unit (Stern 1994) of an assemblage exceeds the time-averaging threshold (Kowalewski 1996) associated with the behavior of interest. Although our finding holds important implications for archaeological interpretations concerning the appearance of behavioral modernity during the Paleolithic, the general effect of time-averaging described here is likely to apply to other periods and regions.

2. The model

We created a spatially explicit agent-based model to address the following research question: to what extent does time-averaging affect the scale of local spatial association in the relative frequency of the most prevalent cultural variant observed in the archaeological landscape? The NetLogo 6.0.2 (Wilensky 1999) source code and full model description following the ODD protocol (Grimm et al. 2010) are freely available at: https://www.comses.net/codebases/cb7a8a57-6e2e-49b1-b785-04b75d7e279/releases/1.0.0/. The R markdown files used to create Figures 2–7 are available as supplementary material (S1.R code). The following is an abridged version of the full model description.

The agents in our model represent hominins (we also refer to these agents as “individuals”) capable of high-fidelity cultural transmission. The population is structured by groups, each occupying a cell on a 20 × 20 grid wrapped around a torus to avoid edge effects. Each group contains \(N = 25\) individuals at the end of each time step. Each time step encompasses the time needed for members of the “experienced” generation to transmit cultural variants to members of the “naïve” generation within each group. Each naïve individual learns its variant of a single selectively neutral cultural trait from a member (or members, depending on the form of cultural transmission represented) of the experienced generation within its group. The model also allows for members of different groups to swap cultural variants via horizontal cultural transmission between groups. The modeled world is abstract and can be imagined as a homogeneous landscape occupied by groups of hominins whose only duties are to traffic in cultural information and then deposit material manifestations of that information (i.e. artifacts) in the cell in which they reside.

The model includes six experimental parameters (Table 1). The transmission of cultural variants from the experienced generation to the naïve generation is central to the dynamics of our model. Because the effects of time-averaging may vary with mechanism of cultural transmission, here we test three widely-recognized forms: unbiased, vertical, and conformist biased cultural transmission (Boyd and Richerson 1985). Under unbiased transmission, each naïve individual randomly selects (with replacement) an experienced member of its group to serve as its teacher. Vertical cultural transmission is modeled in the same way as the genetic transmission of haploid genes during asexual reproduction – each naïve individual inherits its parent’s cultural variant. Under conformist biased cultural transmission, naïve individuals attempt to copy the modal cultural variant of the experienced members of their group. If the majority of a group’s experienced members display “5,” then every member of the naïve generation in that group will attempt to copy “5.” In the event that a group’s experienced generation exhibits more than one mode, each naïve individual in that group chooses randomly from among the modal variants. For instance, assuming “5” and “8” are both modes of a group’s experienced generation, some naïve members of that group may try to copy “5” while others in the same group try to copy “8.”

We assume that intragroup cultural transmission between generations is imperfect. We introduce a second parameter, \(\mu\), to represent the probability (per transmission event) that a naïve individual ultimately adopts a cultural variant other than that which it intends to copy. This in turn requires that we define how error affects the value of the transmitted variant. Here we investigate three different models of copy error, our third experimental parameter. According to a bidirectional single stepwise model of copy error, the naïve individual mistakenly adopts an integer that is one greater than or one less than the target value. For example, a naïve individual who makes an error while attempting to copy the variant “5” will ultimately adopt either “4” or “6” with equal probability. Note that the bidirectional single stepwise model of copy error

| Parameter | Value(s) |
|-----------|----------|
| Local group extinction, \(e\) | 0.01 |
| Intragroup cultural transmission | unbiased, vertical, conformist biased |
| Copy error, \(\mu\) | 0.0001, 0.001, 0.01 |
| Copy error model | bidirectional single stepwise, finite variants, infinite variants |
| Intergroup cultural transmission proportion, \(m\) | 0, 0.05 |
| Geographic extent of intergroup transmission | local, global |
| Duration of assemblage formation, \(d\) | 1, 1000, 10,000 |
allows for “back-innovations,” such that an error results in the same variant displayed prior to the previous copy error. In fact, under the bidirectional single stepwise model, each error results in a “back-innovation” with probability 0.5. We also investigate the effects of time-averaging under a finite variants model of copy error in which the number of possible variants is arbitrarily capped at 100. In this case, each copy error results in the naïve individual being randomly assigned one of the 99 variants that is not its intended target value. The probability of a “back-innovation” (per copy error) in the finite variants model is 1/99. The infinite variants model of copy error eliminates the possibility of back-innovation altogether. With infinite variants, each copying error results in the introduction of a novel cultural variant never before seen during the course of the simulation run.

Our model also includes cultural transmission between members of different groups. The fourth experimental parameter, \( m \), represents the proportion of the population that engages in intergroup cultural transmission. At the start of each simulated time step, 400\( N \text{m}/2 \) pairs of individuals swap cultural variants across group boundaries. The fifth experimental parameter concerns the spatial extent over which horizontal intergroup cultural transmission takes place. With local intergroup cultural transmission, each of 250 randomly chosen agents randomly chooses its intergroup transmission “partner” from among the members of one of the eight groups immediately adjacent to its group. With global intergroup cultural transmission, each of 250 randomly chosen agents randomly chooses its intergroup transmission partner from a group other than its own without respect to the other group’s location on the lattice. Just as gene flow increases genetic diversity within groups and decreases genetic diversity among groups, intergroup cultural transmission increases cultural diversity within groups while decreasing cultural diversity among groups. Intergroup cultural transmission provides an additional mechanism by which cultural variants can spread over space through time.

The sixth experimental parameter is the duration of assemblage formation, \( d \). \( d \) is the number of time steps during which agents deposit cultural variants into artificial archaeological assemblages. Please note that \( d = 1 \) represents a population and \( d > 1 \) a time-averaged assemblage.

Some have argued that the stochastic loss of local groups of hunter-gatherers may have been fairly common during the Middle Paleolithic (Hublin and Roebroeks 2009; Stiner and Kuhn 2006). Middle Pleistocene hominins, especially Neanderthals, probably lived quite high on the trophic pyramid (Hockett and Haws 2005; Richards et al. 2001; Stiner 2002). This is important because the small population densities exhibited by other social carnivores make them more vulnerable to environmental instability and other stochastic events (Diamond 1984; Mckinney 1997), increasing their likelihood of suffering so-called local extinctions. We incorporate this in our model with an additional parameter: \( e \) is the probability per group per time step of a local extinction. Because the purpose of our model is to investigate how time-averaging affects the spatial scale of cultural similarity, we hold constant the probability of local extinction (\( e = 0.01 \)) despite the fact that the effect of \( d \) is at least partly a function of \( e \). Preliminary data collected from simulations without intergroup cultural transmission show that \( d \) has no effect on the spatial scale of cultural similarity when \( e = 0 \), regardless of the value of \( \mu \) or the type of intragroup cultural transmission. We suspect that very high rates of local extinction could also mitigate the effects of \( d \) on the spatial scale of cultural similarity.

To be clear, we do not mean to imply that \( e = 0.01 \) is representative of the actual probability of local group extinction during the MSA in Africa, or during any other period and region. In fact, we are unaware of any empirical data with the temporal and spatial resolution required to confidently estimate \( e \) for the Middle Stone Age in Africa. Here we hold \( e \) constant at 0.01 merely for the pragmatic reason that doing so provides an opportunity to investigate whether time-averaging can affect the spatial scale of cultural similarity in archaeological assemblages. Although we expect the effect of time-averaging to be weaker, and possibly nonexistent, under very low or very high \( e \), we leave the task of identifying those threshold values to future experiments.

### 2.1. Process overview

Each cell in the 20 \( \times \) 20 grid is initialized with \( N = 25 \) individuals at the start of a simulation run. Because each individual in the population starts with a unique cultural variant, the richness of cultural variants in the initial population is equal to the size of the population, or 400\( N \). Each time step represents the time needed for every naïve individual to adopt a cultural variant through cultural transmission. Seven processes occur during every simulated time step in the following order: intergroup cultural transmission, local extinction, local recolonization, aging, creation of a naïve generation, intragroup cultural transmission from experienced to naïve generation, and culling the experienced generation. After 10,000 time steps, an eighth process, cultural variant deposition, occurs at the end of every subsequent time step. 10,000 time steps are sufficient for the richness of cultural variants in the population to reach equilibrium for the parameter values investigated (Figures S1 and S2). Brief descriptions of the eight processes follow.
2.1.1. Intergroup cultural transmission

At the start of each time step 400Nm/2 pairs of agents swap cultural variants across group boundaries at either the local or global scale. Note that intergroup cultural transmission is not susceptible to copy error.

2.1.2. Local extinction

Each group (i.e. each cell) suffers a local extinction event with probability \( e = 0.01 \). Each local extinction event entails the immediate removal of the members of the ill-fated group, rendering the cell vacant.

2.1.3. Local recolonization

In order to fill a cell made vacant by a local extinction event, a group is chosen at random from the subset of cells in the empty cell’s Moore neighborhood (the surrounding eight cells) that contain at least 2 members. The randomly chosen neighboring cell sends half of its current members (rounded up) to serve as the experienced generation of the vacated cell. Note that it is possible for a group to experience multiple fission events during the course of a single time step. One could model recolonization differently, but we feel local recolonization serves as a better approximation of how forager groups expand into available space.

2.1.4. Aging

All individuals who survive local extinction increment their age from “0” to “1.” Any individual with an age of 1 is considered a member of the “experienced” generation who can possibly pass on its variant to a naïve member (or members) of its group.

2.1.5. Creation of the naïve generation

Each cell receives \( N = 25 \) newly created, naïve individuals. Naïve individuals are marked by age = 0. The process by which naïve individuals are created in each cell depends on the intragroup cultural transmission mechanism. With unbiased or conformist biased transmission, where variants are passed obliquely rather than vertically, each cell – regardless of the number of experienced individuals – simply creates \( N \) naïve individuals, each of which is assigned a cultural variant of “0.” The cultural variant of “0” is later replaced by a variant acquired via either unbiased or conformist biased intragroup cultural transmission.

The procedure is different under vertical cultural transmission, whereby every naïve individual attempts to learn its parent’s cultural variant. In the case of vertical cultural transmission, the experienced individuals in each cell take turns producing (and teaching) offspring until the cell contains \( N \) naïve individuals. This means that if \( N \) is a multiple of the number of experienced individuals in a cell, then every member of the experienced generation in that cell will teach the same number of naïve individuals. For example, if there are 5 experienced individuals in a cell, then each will teach 5 offspring, for a total of 25 (5 * 5 = \( N = 25 \)). Likewise, when there are 25 experienced individuals in a cell, then every experienced member of that cell will teach just one offspring (25 * 1 = \( N = 25 \)). However, in cases where \( N \) is not a multiple of the number of experienced individuals in a cell, then at least one of the experienced individuals will teach exactly one more naïve individual than some of the other experienced individuals in its group. For example, if there are 12 experienced individuals left in a cell as the result of a recolonization event, then 11 of them will each teach 2 offspring while one of them will teach 3 offspring ((11 * 2) + (1 * 3) = \( N = 25 \)). Note that compared to unbiased cultural transmission, vertical cultural transmission reduces variance in number of naïve individuals taught among members of the experienced generation within each cell, thereby dampening the effects of drift in intragroup cultural transmission.

2.1.6. Intragroup cultural transmission from experienced to naïve generation

A naïve individual can learn only from the experienced members of its group (i.e. there is no horizontal intragroup cultural transmission). The model includes three different cultural transmission mechanisms: unbiased, vertical, and conformist biased transmission. All three forms of intragroup cultural transmission are susceptible to copy error. Copy error occurs with probability \( \mu \) per transmission event. The variant that results from error is a function of the copy error model. We investigate the effects of \( \mu \) under three different copy error models: bidirectional single stepwise, finite variants, and infinite variants.

2.1.7. Culling the experienced generation

Once cultural transmission is complete and all naïve individuals have adopted a cultural variant, the experienced generation is culled, leaving a population that is the same size (namely, 400N) as the one that was present at the start of the time step.

2.1.8. Cultural variant deposition

Starting with time step 10,001, every individual alive at the end of the time step deposits its cultural variant to the assemblage of the cell it occupies. The growth of the artificial assemblage in each cell mimics the growth of an archaeological assemblage at a particular “locality.” The accumulation of cultural variants across the entire 20 x 20 grid of cells represents the formation of an “archaeological landscape” (Ebert 1992). Access to the assemblage in each cell within the archaeological landscape allows one to investigate how time-averaging affects local spatial association in the most prevalent cultural variant at multiple spatial scales between “locality” and “landscape.” Given that the maximum
value of \( d \) investigated here is 10,000, each simulation ends after 20,000 time steps.

### 2.2. Data collection

We collect data from 20 unique simulation runs at each combination of experimental parameter values. Given \( d \) values of 1, 1000, and 10,000, we collect data after 10,001, 11,000, and 20,000 time steps of each run. We employ Local Moran’s \( I (l) \), a local indicator of spatial association (Anselin 1995), to characterize the spatial scale of cultural similarity in the relative frequency of the most prevalent cultural variant. The most prevalent cultural variant is defined as the variant (i.e. integer) that occurs most frequently in the archaeological landscape (ties broken randomly). For each cell \( i \), where \( i = 1, 2, 3, \ldots n \), \( I_i \) measures the degree of similarity between \( x_i \), the relative frequency of the most prevalent cultural variant in cell \( i \), and the relative frequency of the most prevalent cultural variant displayed in the \( j \) “neighboring” cells, \( x_j \). \( I_i \) is calculated as follows:

\[
I_i = \frac{z_i}{m_2} \sum_j w_{ij}z_j, \tag{1}
\]

where \( z_i \) and \( z_j \) are deviations of \( x_i \) and \( x_j \), from the mean of the non-categorical variable \( X \), \( m_2 = \sum (z_i^2/n) \), and \( w_{ij} \) is a binary spatial weight in which each pair of neighbors \( i \) and \( j \) is designated with a 1 and each non-neighbor pair is assigned a 0 (Anselin 1995; Sokal, Oden, and Thomson 1998).\(^1\) In this study, the size of the square neighborhood surrounding the \( i \)th cell is defined by the spatial lag. Spatial lags of 1, 2, 3, 4, 5 and 10 correspond to 8, 24, 48, 80, 120, and 399 neighboring cells, respectively. \( I_i \) scores can be standardized under the hypothesis of total randomization by subtracting an expected value \( E[I] \) from the raw \( I_i \) score (Equation 1) and then dividing the difference by the standard deviation of the raw \( I_i \) scores (Anselin 1995; Sokal, Oden, and Thomson 1998). The expected value is given by:

\[
E[I_i] = -w_i/(n-1) \tag{2}
\]

where \( w_i = \Sigma_{j\neq i} w_{ij} \) is the sum of the \( i \)th row’s elements (the number of neighbors in our case) (Sokal, Oden, and Thomson 1998). Standardized \( I_i \) scores are positive when \( x_i \) and \( x_j \) are relatively high or relatively low values. When \( x_i \) and \( x_j \) are dissimilar, standardized \( I_i \) is negative. Values approach zero when there is no spatial association between \( x_i \) and \( x_j \). The sign of a standardized \( I_i \) score does not reflect whether the spatial neighborhood is marked by relatively high or relatively low values compared to the rest of the landscape, but rather only the degree to which the set of values \( (x_i) \) displayed by neighboring cells are similar to each other and to \( x_i \) (resulting in a positive \( I_i \)) or dissimilar to \( x_i \) (resulting in a negative \( I_i \)). We used the “spdep” package (Bivand and Piras 2015; Bivand, Hauke, and Kossowski 2013) in R to calculate \( I_i \).

In our case, \( x_i \) is the \( i \)th cell’s relative frequency of the most prevalent cultural variant, as defined above, and \( x_j \) refers to the relative frequencies of the same (i.e. the most prevalent) cultural variant in \( j \) neighboring cells. For each value of \( d \) we calculate standardized \( I_i \) (hereafter, simply \( I_i \)) for each cell at six spatial lags. Because we wish to characterize the scale of local spatial association displayed by an archaeological landscape, we investigate mean \( I_i \) at each spatial lag. Note that the mean of the standardized local \( I_i \) values for a given spatial lag is not equivalent to global Moran’s \( I \).

For the purpose of illustration, Figure 1 provides the \( I_i \) results for a single simulation run at two different values of \( d \). Note that the integer value of the most prevalent variant for \( d = 1 \) is not necessarily the same as the integer value of the most prevalent variant for \( d = 10,000 \).

---

**Figure 1.** Relative frequencies of the most prevalent variant (blue) and the corresponding standardized \( I_i \) scores for spatial lags of 1, 2, 3, and 4 (red). All data come from a single simulation run \( (\epsilon = 0.01, \mu = 0.001, m = 0) \), unbiased cultural transmission, finite cultural variants copy error model) at different values of \( d \); \( d = 1 \) (top row) and \( d = 10,000 \) (bottom row).
3. Results

The experimental parameter of primary interest is \( d \), the duration of assemblage formation. We calculate Local Moran’s \( I_i \) at six different spatial lags on assemblages that form over \( d = 1, 1,000, \) and \( 10,000 \) time steps. Given the possibility that the effect of \( d \) is a function of other closely related parameters, we also vary the probability of copy error during intragroup cultural transmission \((\mu)\), the type of copy error model (bidirectional single stepwise, finite variants, or infinite variants), the mechanism of intragroup cultural transmission from the experienced generation to the naive generation (unbiased, vertical, or conformist biased), and the spatial extent (local or global) of intergroup cultural transmission. We present our results in two parts. First, we present data from simulations that lack intergroup cultural transmission \((m = 0)\). Second, we present data from simulations in which 250 pairs of agents engage in intergroup cultural transmission at the start of each time step \((m = 0.05)\).

3.1. The effect of \( d \) on the spatial scale of cultural similarity in the absence of intergroup cultural transmission \((m = 0)\)

In the absence of intergroup cultural transmission, \( d \) has a positive effect on the spatial scale of cultural similarity under most of the conditions we investigated (Figures 2–4). The effect of \( d \) on the spatial scale of cultural similarity is perhaps best described with an analogy to the cross-section of a wave. As \( d \) increases

![Figure 2. The effect of time-averaging on the spatial scale of cultural similarity under the bidirectional single stepwise model of copy error. Assemblage duration increases from left to right: \( d = 1, \) \( d = 1,000, \) and \( d = 10,000. \) Intrigroup cultural transmission varies by row: unbiased transmission (top), vertical transmission (middle), and conformist biased transmission (bottom). Each data point represents the mean ± 1 standard deviation of mean \( I_i \) from 20 unique simulations. Solid lines connect the means. Color represents different levels of copy error: \( \mu = 0.0001 \) (blue); \( \mu = 0.001 \) (green), and \( \mu = 0.01 \) (red). \( m = 0 \) for all runs.](image-url)
(moving from the left panel to the right panel in each row), the highest point, or “crest,” of the wave moves from shorter spatial lags to longer spatial lags. Consider just the top row of Figure 2, which provides results collected under the bidirectional single stepwise copy error model and unbiased cultural transmission. At $d = 1$, mean $I_i$ values are relatively close to 0 regardless of spatial lag. The highest mean $I_i$ values are found at relatively short spatial lags (lag of 1, 2, or 3, depending on the value of $\mu$). Increasing $d$ from 1 to 10,000 (i.e. moving from the left panel to the right panel within a row) has two important effects on mean $I_i$ values. First, the mean $I_i$ values increase for spatial lags less than 10 but decrease at a spatial lag of 10. Following our wave analogy, the decrease in mean $I_i$ values at a spatial lag of 10 is reminiscent of the undertow that accompanies a wave approaching shore. Second, the crest of the wave, i.e. the highest point on the curve connecting the means of mean $I_i$ values, “travels” to longer spatial lags (to lags of 3 or 4, depending on $\mu$) as $d$ increases from 1 to 10,000. In this case, increasing $d$ increases the spatial scale of cultural similarity in the relative frequency of the most prevalent cultural variant.

Figure 2 presents results for the bidirectional single stepwise model of copy error, for which the probability that each copy error results in a “back-innovation” is 0.5. Here, increasing $d$ increases the spatial scale of cultural similarity in the most prevalent cultural variant regardless of whether intragroup cultural transmission is unbiased, vertical, or
conformist biased. Under unbiased cultural transmission, the magnitude of the effect of $d$ on the spatial scale of cultural similarity is greater for lower $\mu$. By contrast, $\mu$ does not influence the effect of $d$ under vertical or conformist biased intragroup cultural transmission.

Figure 3 presents results for the finite variants model of copy error. Given that our finite variants model is capped at 100 unique variants, the likelihood that a copy error results in a "back innovation" is nearly 50 times lower than under the bidirectional single stepwise model of copy error. Figure 3 shows that $d$ does not always affect the spatial scale of cultural similarity in the most prevalent cultural variant. In other words, the effect of $d$ depends in part on the copy error model, the rate of copy error, and the mechanism of intragroup cultural transmission. While increasing $d$ does not increase the spatial scale of cultural similarity when copy errors are quite frequent ($\mu = 0.01$) under either unbiased or vertical intragroup cultural transmission, the effect of $d$ on the spatial scale of cultural similarity is robust to higher values of $\mu$ under conformist biased transmission.

The infinite variants model of copy error further reduces the likelihood of a "back innovation" to 0. Due to computational limitations, we were able to collect data under the infinite variants model for just one value of $\mu$, the slowest ($\mu = 0.0001$). Figure 4

Figure 4. The effect of time-averaging on the spatial scale of cultural similarity under the infinite variants model of copy error. Assemblage duration increases from left to right: $d = 1$, $d = 1,000$, and $d = 10,000$. Intragroup cultural transmission varies by row: unbiased transmission (top), vertical transmission (middle), and conformist biased transmission (bottom). Each data point represents the mean ± 1 standard deviation of mean $I_i$ from 20 unique simulations. Solid lines connect the means. $\mu = 0.0001$ and $m = 0$ for all runs.
shows that the infinite variants results are qualitatively identical to those obtained under the finite variants model for the same value of \( \mu \), and there are good reasons to presume that the infinite variants results would look similar to the finite variant results (Figure 3) at higher values of \( \mu \).

### 3.2. The effect of \( d \) on the spatial scale of cultural similarity in the presence of intergroup cultural transmission (\( m = 0.05 \))

There are a number of ways to represent intergroup cultural transmission in a structured population (e.g. Premo 2012). Because our goal is to develop a general sense of how intergroup cultural transmission might weaken the effect of \( d \) reported above, we simply vary the geographic extent (local or global) within which members of different groups are chosen to participate in intergroup cultural transmission while holding constant copy error (\( \mu = 0.0001 \)) and the proportion of agents engaged in intergroup cultural transmission (\( m = 0.05 \)).

Figure 5 presents results for the bidirectional single stepwise model of copy error with intergroup cultural transmission (\( m = 0.05 \)) at the local (red) and global (green) scale. Assemblage duration increases from left to right: \( d = 1 \), \( d = 1,000 \), and \( d = 10,000 \). Intrigrup cultural transmission varies by row: unbiased transmission (top), vertical transmission (middle), and conformist biased transmission (bottom). Each data point represents the mean ± 1 standard deviation of mean \( I_i \) from 20 unique simulations. Solid lines connect the means. \( \mu = 0.0001 \) for all runs. Note that in some cases, the green data points completely obscure the red data points.
global scale. In other words, for unbiased and vertical cultural transmission \( d \) has no effect on the spatial scale of cultural similarity when the partners involved in intergroup cultural transmission are chosen from groups without respect to the distance that separates them. The picture is different for conformist biased cultural transmission, however. Here, increasing \( d \) increases the spatial scale of cultural similarity in the most prevalent cultural variant regardless of whether the scale of intergroup cultural transmission is local or global.

The same is true under both the finite variants model of copy error (Figure 6) and the infinite variants model of copy error (Figure 7), which reduce the probability that each copy error results in a "back-innovation" to 1/99 and 0, respectively. Under unbiased and vertical cultural transmission, increasing \( d \) increases the spatial scale of cultural similarity under local – but not global – intergroup cultural transmission. Here, again, the positive effect of \( d \) on the spatial scale of cultural similarity is robust to local and global intergroup cultural transmission under conformist biased intragroup cultural transmission.

### 4. Discussion

In the presence of group extinction and local recolonization there are many conditions in which increasing the duration of assemblage formation increases the
scale of local spatial association in the relative frequency of the most prevalent cultural variant. How can one explain this effect?

Local extinctions beget local recolonization, whereby a randomly chosen subset of the members of an adjacent surviving group carry cultural variants into the empty cell and are given the chance to pass them on to the next generation of that cell’s inhabitants. Local extinction followed by local recolonization has the same general effects as gene flow among adjacent groups in a stepping stone model of population structure (Kimura and Weiss 1964): it 1) inhibits differentiation among groups and 2) provides a mechanism by which a variant can spread from the group in which it first appears to other groups within the metapopulation.

In our model, local extinction is stochastic and all cultural variants are selectively neutral. Thus, in much the same way that sampling error is responsible for the rise and fall of the frequency of selectively neutral variants in a finite population (e.g. Neiman 1995; Bentley, Hahn, and Shennan 2004), the stochastic nature of local group extinctions and local recolonization ensures that some selectively neutral cultural variants will come to be more widely distributed across space than others before eventually succumbing to the very same stochastic force and disappearing from the population. Owing to nothing more than the

Figure 7. The effect of time-averaging on the spatial scale of cultural similarity under the infinite variants model of copy error with intergroup cultural transmission ($m = 0.05$) at the local (red) and global (green) scale. Assemblage duration increases from left to right: $d = 1$, $d = 1,000$, and $d = 10,000$. Intragroup cultural transmission varies by row: unbiased transmission (top), vertical transmission (middle), and conformist biased transmission (bottom). Each data point represents the mean ± 1 standard deviation of mean $I_i$ from 20 unique simulations. Solid lines connect the means. $\mu = 0.0001$ for all runs. Note that in some cases, the green data points completely obscure the red data points.
“dumb luck” of being chosen to relocate to an empty cell as part of the recolonization process and then getting adopted by at least one member of the naïve generation in that newly colonized cell, a variant can spread over space as it also increases in frequency in the population (and in the archaeological landscape) through time.

Local extinction and local recolonization not only decrease the equilibrium level of cultural diversity at the level of the metapopulation (Premo and Kuhn 2010), they also structure cultural variation within and between groups in a way that allows for the formation of spatially contiguous “clusters” of groups whose members are culturally similar due to their recent shared history of cultural transmission from experienced members of the same “ancestral” group. The variants displayed within such large clusters of culturally similar groups enjoy safety in numbers – that is to say, they enjoy some protection against the stochastic process of local extinction – relative to variants displayed within smaller clusters of culturally similar groups. Because recolonizers are imported from an adjacent cell, a cell vacated within a large cluster of culturally similar groups is more likely to be colonized by individuals carrying cultural variants that are identical to those displayed by the group that previously inhabited the cell than is a cell vacated within a “cluster” of just 2 or 3 culturally similar groups. Thus, in the presence of local recolonization, the sheer size of a cluster of culturally similar groups can prolong the period during which the cultural variants that distinguish those groups from other (more distantly “related”) groups are deposited into archaeological assemblages. With time, the archaeological assemblages in cells within a cluster of culturally similar groups become more similar to one another relative to the assemblages recovered from two cells chosen randomly from the entire archaeological landscape.

The size of the transient clusters of culturally similar groups changes through time – starting small with a single group, growing to include many groups, and then shrinking and disappearing altogether. The cycle that includes the growth, perseverance (due to safety in numbers), demise, and ultimate disappearance of clusters of culturally similar groups has been documented in similar models in population genetics (e.g. the stepping stone model [Kimura and Weiss 1964]) and economics (e.g. the voter model [Clifford and Sudbury 1973; Cox and Griffeath 1986; Weidlich 1971]). In all of these models, clusters of similar groups expand, persist, and then contract through time. The average length of a cluster’s “lifespan” and the extent of its geographic “footprint” depend upon copy error rate, extinction rate, and the nature of the recolonization rule.

Due to the stochastic nature of local extinction and recolonization, clusters of culturally similar groups not only change size through time but also “drift” (not to be confused with random genetic drift) over space. Even as one side of the cluster is nibbled away by local extinction and recolonization the other side could be expanding as a result of the same stochastic processes. Just as a localized storm system that lingers over a region for days drops precipitation on an area much larger than the parcel of land that receives rain during any single hour during that period, clusters of culturally similar groups deposit variants in the assemblages of a greater number of cells than were occupied at any single point in time by those groups. This helps explain those cases in which the scale of local spatial association in the relative frequency of the most prevalent cultural variant increases with $d$. This finding has an important implication for archaeological inference. In the same sense that it might be difficult to infer the size of the active storm cloud during any one-hour period from the spatial extent of the area of land that received at least some rain over the course of a 10-day storm, it might prove difficult to infer the spatial scale of the clusters of culturally similar groups in any past population given data collected from heavily time-averaged assemblages.

The discussion up to this point applies to conditions in which intergroup cultural transmission is either absent ($m = 0$) or occurs at the local scale thereby mimicking the effects of migration in a two-dimensional stepping stone model. It is instructive to address both why time-averaging does not affect the spatial scale of cultural similarity for unbiased and vertical transmission in the presence of global scale intergroup cultural transmission and why the effect is robust to global scale intergroup cultural transmission under conformist biased cultural transmission. As stated above, intergroup cultural transmission provides an additional mechanism by which cultural variants can spread to other groups through time. Global scale intergroup cultural transmission can short-circuit the engine (local extinction and local recolonization) that drives the recursive cycle that includes the initial expansion, temporary insulation, and ultimate demise of regional clusters of culturally similar groups. Indeed, we see that global scale intergroup transmission swamps the effects of the local extinction and recolonization under unbiased and vertical cultural transmission.

But our results clearly show a different picture under conformist biased transmission. When intergroup cultural transmission is conformist, the effect of time-averaging on the spatial scale of cultural similarity is robust to global scale intergroup cultural transmission (at least for $m = 0.05$) (Figures 5–7). To better understand this difference, first characterize the strength of the effect of intergroup cultural transmission as the proportion of variants introduced from outside the group that are passed on to the naïve generation during
subsequent intragroup cultural transmission. Under vertical transmission, 100% of the variants introduced to a group via intergroup cultural transmission are passed to the naïve generation. Under unbiased cultural transmission, the proportion of variants introduced by intergroup cultural transmission that are passed to the naïve generation is subject to sampling error in a finite-sized group. However, because we know that the expected proportion will be less than 100% for any finite-sized group, the effect of intergroup cultural transmission is weaker under unbiased transmission than vertical transmission in models that do not assume infinitely large groups. The effect of intragroup cultural transmission is weakest under conformist biased transmission, because none of the novel variants introduced via intergroup cultural transmission are transmitted to the naïve generation. Conformist biased intragroup cultural transmission negates the effects of intergroup cultural transmission by quickly “erasing” the few unique cultural variants acquired from the outside. Because conformist biased transmission essentially makes the variants introduced by intergroup cultural transmission invisible to the naïve generation, even global scale intergroup transmission cannot mitigate the effect of time-averaging on the spatial scale of cultural similarity in the most prevalent cultural variant. The fact that conformity can render global intergroup cultural transmission toothless in this context is an important lesson for archaeologists interested in dealing with the effects of time-averaging on the spatial scale of cultural similarity, especially if conformist biased cultural transmission is as prevalent in human societies as has been argued (e.g. Henrich 2001; Henrich and Boyd 1998; Henrich and McElreath 2003).

Of course, the usual disclaimer applies: the results should only be interpreted in the context of our model assumptions. Our spatial analysis considers the relative frequencies of only the most prevalent cultural variant in the archaeological landscape. This was a pragmatic choice. The variant that appears most frequently in the archaeological landscape is also the one that potentially could show the largest scale of local spatial association. In short, we chose to focus on the kind of cultural data most likely to inform on the issue of whether d affects the scale of local spatial association in culturally transmitted traits. Obviously, concentrating on the most prevalent cultural variant of a single trait does not allow us to address related issues such as diversity at multiple cultural traits. Our approach is perhaps most reminiscent of cases in which archaeologists treat the relative frequency of a particular variant or tool type as diagnostic of a larger cultural tradition.

Second, our model considers local recolonization only. When a group suffers local extinction, all of the recolonizers are derived from one of its eight neighboring cells. Of course, there are other ways to model recolonization: recolonizers could be drawn from multiple cells within the Moore neighborhood, they could all come from one cell chosen randomly from the entire landscape, or they could be chosen randomly from the metapopulation at large. We wish to stress that the details of recolonization matter. For example, non-local forms of recolonization would drastically decrease the likelihood that a newly recolonized group would be more similar to any of its immediate neighbors than to any other group in the metapopulation. Under such conditions, one would not expect mean I_i to vary with spatial lag, holding d and e constant. Nor would one expect d to affect mean I_i, holding spatial lag and e constant.

Third, the cultural variants in our model are selectively neutral – variants do not affect an individual’s ability to survive a local extinction event or to improve its likelihood of colonizing an empty cell. Including natural selection could significantly alter the results discussed here, although exactly how would depend in part on the nature of the fitness landscape of the cultural trait. For example, if the fitness landscape for tool form had just one global optimum, then e and d would have little effect on the scale of local spatial association in the most prevalent cultural variant in the assemblage. In that case, recolonizers would look similar to the members of the groups they replace not due to shared “cultural ancestry” but due to the fact that all groups are tracking the same environmental signal (or, converging on the same fitness peak) with their technology. In the case of a rugged fitness landscape, in which the fitness effects of tool form vary by region (i.e. the form that works the best in one region does not work well in the others), one would expect e and d to have little effect on the scale of local spatial association in the most prevalent cultural variant for a different reason. Although recolonizers that come from cells in the “valley” between two local fitness “peaks” might initially display a different tool form from the previous and better adapted occupants of the cell, they will soon adopt a similar form if they have recourse to individual learning or they will be replaced by a neighboring group that does. In both cases, the size of the clusters of culturally similar groups is determined primarily by the topology of the fitness landscape of tool form, not by e and d.

5. Conclusion

Regional cultural differentiation is a characteristic of a population – a set of people alive at the same time. For better or for worse, Paleolithic archaeologists have access to assemblages of artifacts, not past populations. Because of the importance of the notion that regional cultural differentiation might be a sign of behavioral modernity, we set out to investigate to what extent time-averaging might distort the view of the spatial scale of regional cultural differentiation afforded by
archaeological assemblages. Although our primary research interest is in the Paleolithic, we would argue that our basic findings apply more generally to any time-averaged assemblage.

Our results suggest that time-averaged assemblages may not necessarily reflect the true spatial scale of the “culture regions” exhibited by any of the past populations that deposited artifacts to the archaeological record. Simulations show that time-averaging can increase the scale of local spatial association among the relative frequencies of the most prevalent cultural variant. As a spatial cluster of culturally similar groups expands, “drifts” about a landscape, and ultimately contracts and disappears according to the vagaries of local extinction and recolonization, its members deposit cultural material over a larger area than that occupied by the cluster of similar groups at any single point in time. Thus, one should expect the signal observed in the assemblage to depart from that observed in any population whenever the minimum archaeological-stratigraphic unit (Sten 1994) of an assemblage exceeds the time-averaging threshold associated with the local extinction and recolonization of hominin territory. This should bring to mind the difficulty of inferring the size of an active rain cloud at any given hour from a map of total precipitation per hectare over a 10-day period.

Despite the abstract nature of our model (or perhaps, because of it), the results point to an interesting question: Might time-averaging be at least partly responsible for the apparent difference between the spatial scale of regional cultural differentiation during the MSA and LSA? As is often the case, this question points to yet another: are MSA assemblages generally more time-averaged than LSA assemblages? While every assemblage has its own depositional history and thus it is impossible for any single statement to apply to each and every MSA and LSA assemblage, Perreault (2012, 2) argues that older archaeological assemblages are generally associated with longer intervals of time. If he is correct, then the appearance of a greater number of smaller, regionally distinctive “cultures” in the archaeological record of the LSA – an archaeological signal traditionally interpreted as a unique by-product of behavioral modernity – could be explained in part by the fact that LSA assemblages are on average less time-averaged than their MSA counterparts. In our opinion, the results of our heuristic modeling exercise do not provide “the answer” so much as additional motivation to address such interesting and previously underappreciated questions empirically.

Acknowledgements

We thank Mark Lake and three anonymous reviewers for their insightful and helpful comments on the penultimate version of this paper.

Disclosure statement

No potential conflict of interest was reported by the authors. The underlying research materials for this article can be accessed at: https://www.comses.net/codebases/cb7a8a57-6e2e-49b1-b785-04b75d7e6279/releases/1.0.0.

Notes on contributors

Galen Miller-Atkins received his B.A. from Ohio University in 2014 and his M.A. from Washington State University in 2016. He has conducted Paleolithic fieldwork in Georgia and South Africa.

L. S. Premo is Associate Professor of Anthropology at Washington State University in Pullman, WA, USA and an affiliated researcher in the Department of Human Evolution at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. He received his B.A. from Colorado State University in 1999 and his Ph.D. from University of Arizona in 2006. He uses computer simulation to study cultural evolution. His research has been published in Proceedings of the National Academy of Sciences, Current Anthropology, and American Antiquity.

References

Anselin, L. 1995. “Local Indicators of Spatial Association — LISA.” Geographical Analysis 27: 93–115. doi:10.1111/j.1538-4632.1995.tb00338.x.

Bailey, G. N. 1983. “Concepts of Time in Quaternary Prehistory.” Annual Review of Anthropology 12: 165–192. doi:10.1146/annurev.an.12.100183.001121.

Bailey, G. N. 2007. “Time Perspectives, Palimpsests and the Archaeology of Time.” Journal of Anthropological Archaeology 26: 198–223. doi:10.1016/j.jaa.2006.08.002.

Bentley, R. A., M. W. Hahn, and S. J. Shennan. 2004. “Random Drift and Culture Change.” Proceedings of the Royal Society B: Biological Sciences 271: 1443–1450. doi:10.1098/rspb.2004.2746.

Binford, L. R. 1981. “Behavioral Archaeology and the “Pomei Premise’.” Journal of Anthropological Research 37: 195–208.

Bivand, R. S., J. Hauke, and T. Kossowski. 2013. “Computing the Jacobian in Gaussian Spatial Autoregressive Models: An Illustrated Comparison of Available Methods.” Geographical Analysis 45 (2): 150–179.

Bivand, R. S., and G. Piras. 2015. “Comparing Implementations of Estimation Methods for Spatial Econometrics.” Journal of Statistical Software 63 (18): 1–36.

Boyd, R., and P. J. Richerson. 1985. Culture and the Evolutionary Process. Chicago: University of Chicago Press.

Note

1. Note that the spatial weights matrix used to calculate \( I_i \) need not be binary. We also calculated \( I_i \) values with row-standardized spatial weights, which are obtained by dividing the binary value described above (i.e., 1 for neighbors of \( i \) or 0 for non-neighbors of \( i \)) by the number of \( i \)'s neighbors (again, not including \( i \)). The results are qualitatively the same as those obtained with binary spatial weights.
