Simulating cultural transmission: preliminary results and their implications for the study of the formal variability of material culture in the Central Balkan Neolithic

Marko Porčič, Miloš Nešić
Department of Archaeology, Faculty of Philosophy, University of Belgrade, RS
mporcic@f.bg.ac.rs

ABSTRACT – In this paper, we adopt the theoretical framework of evolutionary archaeology in order to model and simulate cultural transmission between hypothetical Neolithic sites in Balkans. We simulate neutral cultural transmission in order to compare the simulation results with empirically observed patterns of material culture variability such as traditional archaeological cultures. Our preliminary results show that a series of random local interactions can result in spatial groupings of typologically similar assemblages that correspond to the spatial distributions of traditional archaeological cultures, even in the absence of any other ‘external’ factor such as an overarching regional political structure or shared collective identity.


KEY WORDS – evolutionary archaeology; cultural transmission; archaeological culture; Neolithic; Balkans; simulation

Introduction

In Neolithic research, as in any other research, great effort is invested in looking for patterns. However, the search for patterns is only the first step. The final aim is to account for these patterns in terms of the historical and anthropological dynamics that produced them. The standard procedure in science is to propose an explanation (hypothesis), derive the empirical implications (expectations) from this hypothesis, and then compare these expectations to the actual empirical situation. Our hypotheses about the past are often about complex processes, and archaeological data are equally complex, so it is difficult to explore the implications of our hypotheses without the aid of some formal method. One possible approach to this problem involves attempts to ‘recreate’ the past by constructing a model related to some aspects of the past and then exploring the behaviour of the model and its output by computer simulation (Lake 2014). For example, geneticists have simulated genetic effects related to different scenarios of Neolithisation (e.g., François et al. 2010) in order to see which scenario would produce contemporary genetic patterns in space as revealed by Cavalli-Sforza’s seminal research (Cavalli-Sforza 2001). Archaeologists have simulated the demographic dynamics of the spread of the Neolithic in order to account for patterns related to radiocarbon and settlement evidence (e.g., Fort et al. 2012, Lemmen et al. 2011).

From an epistemological perspective, the simulation approach enables archaeologists to do what is gene-
rally very difficult in social sciences and utterly impossible for historical disciplines: to approximate the experimental method by ‘repeating’ different versions of history and observing the outcomes (Grüne-Yanoff, Weirich 2010; Lake 2014).

However, what makes simulation studies possible is the theoretical framework. For example, genetic simulation models are constructed using concepts and principles of population genetics theory, and demographic simulations are based on demographic theory, which provides the conceptual framework and mathematical models of population dynamics. But what about the formal variability of material culture, which is the traditional domain of archaeology? The issue of style has been the central issue of the traditional culture-historical approach, which is still the dominant school of thought in many academic communities, especially in Southeastern Europe. In the traditional approach, variability in form was divided into entities called archaeological cultures; these entities were both patterns and explanations at the same time, because they were based on an essentialist view of archaeological cultures as direct reflections of collective identities: ethnic, linguistic, political (or even racial) ( Hodder 1982:2–12; Shennan 1994). But regardless of the fact that such traditional culture-historical explanations are outdated, the problem remains: how can we account for the formal variability of material culture in time and space (for the most recent and thoughtful discussion of this problem, see papers in Roberts and Vander Linden (2011))? Just as there are genetic and demographic patterns, there are also patterns of formal variability of material culture. So, is there a theory, other than traditional culture-historical theory, that can provide a suitable framework for translating hypotheses about the workings of past societies into formal models whose properties and implications can be investigated by means of computer simulations and then compared to patterns of material culture variability observed in the archaeological record?

Evolutionary theory of culture – a new framework for an old problem

The evolutionary theory of culture or cultural transmission theory, a relatively recent development in the history of archaeological thought, is a paradigm that provides the intellectual and analytical tools to translate the patterns of formal variation of material culture in time and space into meaningful and anthropologically relevant statements about the past (Lipo 2001; O’Brien, Lyman 2000; 2003; Shennan 2002; 2011). It is based on an evolutionary theory of culture that views culture as an evolutionary process (Boyd, Richerson 1985; Cavalli-Sforza, Feldman 1981; Mesoudi 2011; Mesoudi et al. 2006; Richerson, Boyd 2005; Shennan 2002; 2011). The evolutionary view of culture allows different classes of material culture to be treated as different hereditary systems, thus enabling the analyst to infer the nature and trajectory of cultural transmission and its underlying behavioural and social basis. This kind of analysis has the potential to tackle the most intriguing questions about the anthropological and historical reality that lies behind the archaeological record (e.g., Bentley, Shennan 2003; Bettinger, Eerkens 1999; Gray, Atkinson 2003; Gray, Jordan 2000; Jordan, Shennan 2003; Lipo 2001; Lipo et al. 1997; Lycett 2007; Neiman 1995; Tehrani, Collard 2002; Tehrani et al. 2010).

In this theory, culture is conceptualised as a population phenomenon: each domain of culture can be characterised by a frequency distribution of traits. These traits are culturally transmitted in a process that is analogous in some degree to the transfer of genes (Henrich et al. 2008). Changes in frequencies of cultural traits are governed by the forces of cultural evolution, such as drift or various forms of selection (Boyd, Richerson 1985; Richerson, Boyd 2005). This means that attributes, types and assemblages can be thought of as the results of various cultural transmission processes operating at different scales (cf. Clarke 1968).

A growing number of simulation studies investigate the properties and implications of different transmission models in time and space (e.g., Crema et al. 2014; Lipo et al. 1997; Premo, Scholnick 2011; White 2013). These studies are of great theoretical and methodological importance. In this paper, we aim to make a link between abstract models from cultural transmission theory and the specific context of the Central Balkan Neolithic.

Models of cultural transmission

There are several basic transmission models in evolutionary theory of culture (Boyd, Richerson 1985; Richerson, Boyd 2005; Shennan 2002). These models tell us how a frequency of a trait will behave in time if it is transmitted in a certain way. One of the most important models in evolutionary archaeology is the neutral model of cultural transmission (Bentley, Shennan 2003; Eerkens, Lipo 2005; Kohler et al. 2004; Neiman 1995; Premo, Scholnick 2011;
Simulating cultural transmission: preliminary results and their implications for the study of the formal variability ...  

Shennan, Wilkinson 2001; Steele et al. 2010). The neutral model assumes a random copying of cultural traits in a population. This can be illustrated, for example, by a group of potters randomly copying patterns of vessel ornamentation from one another (Fig. 1). Initially, each potter decorates a bowl with a distinctive motif. After some time, equal to the average use-life of bowls, each vessel from the set is broken and deposited, and each potter creates a new vessel by randomly deciding how to decorate each new bowl. Potters can choose a pattern of ornamentation for each new bowl from the previous generation of bowls or they can introduce a completely new motif (i.e. new to this community of potters) either by intention or by error of perception with an associated probability (μ), which is analogous to the mutation concept in biology. Following Neiman, the probability of mutation (μ) can be broken down into two components μ = ν + m (Neiman 1995.17), where m is the probability that the new variant will be introduced from another community and ν is the probability that a completely new variant will be introduced.

It should be emphasised that in cultural contexts the neutral model does not have to be necessarily interpreted in such a way that individuals copy traits (e.g., pottery decoration motifs) completely at random. Individual choices can be and often are idiosyncratic and have a meaning for the individual making the choice, but as long as the aggregate result of these individual choices is such that the probability of copying for each trait depends only on its current frequency in the population (e.g., corresponding to the simple interpretation of the model presented above), the transmission process is effectively random (Shennan 2011.1073).

In general, if the transmission is not neutral, it is biased in some manner. There are various forms of biased transmission. For example, the conformist bias is one such model; it can be illustrated by many examples where people show a tendency to conform by choosing the most common cultural trait in the population, ranging from religious beliefs to choosing a hair style or clothing style. In conformist transmission, there is an additional probability (degree of conformism) that an entity will copy the most frequent variant in the population. However, this preliminary report is limited to the simulation of the neutral model.

Research aims, questions and hypotheses

The general idea of this paper is to illustrate how cultural transmission models for the Neolithic of the Central Balkans can be translated into computer simulations. The aim is to make a simulation of cultural transmission that enables the analyst to answer theoretical, methodological and empirical questions related to the patterns of formal variability of material culture in the Central Balkans. This project is ongoing, and in this paper we present only preliminary solutions and preliminary results. Due to its incompleteness and lack of suitable empirical data for rigorous testing, this paper should be viewed only as an illustration of the potential of the simulation approach grounded in the evolutionary theory of culture. Specifically, we address two questions in this study:

1. Can simulations produce patterns that resemble the empirical patterns of formal variability of material culture in the Central Balkan Neolithic? Our hypothesis is that the neutral model of transmission coupled with the specifics of Balkan topography is sufficient to produce patterns of typological similarity observed in the archaeological record. If correct, this hypothesis would imply that there is no need to invoke collective identity explanations (such as ethnic, linguistic or political) to explain the observed empirical patterns in the distribution of pottery styles.

2. Should we expect the geographical distance to be strongly correlated with typological distance if the probability of interactions between sites is determined by their spatial distance? A recent empirical study of Iroquian pottery demonstrated that geographical distance had little effect on pottery similarity...
John P. Hart (2012) found that the geographical distance explained a relatively small amount of inter-assemblage typological variance. We use the simulation to see whether the neutral transmission, which is often used as underlying model for the evolution of style (Cochrane 2001; Dunnell 1978; Lipo et al. 1997; Neiman 1995), can produce such a result.

**The simulation**

**Geographical set-up**
The starting point for all simulations is a set of virtual sites generated on a topographical map of Central and Western Balkans (Fig. 2). We chose computer-generated site locations instead of real site locations because of the unequal and biased research history in Balkans, which makes current data about the spatial distribution of Neolithic sites unreliable. Site locations are chosen according to the most general Neolithic settlement criteria (e.g., low altitude, flat terrain, proximity to water). In total, 100 sites were generated in this simulation. When site locations are chosen, the computer calculates the least cost path (LCP) distance between each pair of sites and stores these distances in the LCP distance matrix.

**Sites properties and virtual material culture assemblage**
The material culture assemblage for each site is modelled as a vector of integers, each element of the vector representing a certain variant of that particular material culture class. These vector elements will be referred to as entities. Entities may carry different variants (i.e., have different integer values); this can be interpreted as, for example, different ceramic bowl shapes or different decorative motifs. The size of the material culture assemblage (N) is 100 entities for each site. The variants which entities carry are culturally transmitted.

**Modeling cultural transmission**
For the purpose of this preliminary report, we simulate only the neutral model of cultural transmission. The simulation of the neutral model is based on the standard algorithm presented in the literature (Bentley et al. 2004; 2007; Bentley, Shennan 2003; Crema et al. 2014; Neiman 1995).

For each site, in each iteration, and for each entity, the computer generates a random number between 0 and 1, and based on the generated value, the entity chooses between three options:

1. If a random number falls between $\mu$ and 1, the entity will randomly copy a variant from another entity from its own site, including itself, with the probability of a particular variant being copied equal to its current relative frequency in the assemblage.

2. If the random number falls between 0 to $\nu$, the entity will generate a completely new (to the entire simulation universe) variant.

3. If the random number falls between $\nu$ and $\mu$ (being equal to $m = \mu - \nu$, see below) an entity will copy a variant from another site. The probability of each site being chosen as the source of the new variant is proportional to its LCP distance from the focal site. When the site to be copied from is chosen, a variant to be copied is chosen randomly from the cultural assemblage of that site.

**Simulation procedure and output**
The computer records the assemblage structure (i.e., the relative frequency of variants) for each site for each iteration. Each iteration corresponds to one cultural generation, because if entities are interpreted as ceramic bowls, then one simulation time step is
equal to the bowl’s use-life, assuming that a broken bowl needs to be replaced with a new bowl and that the making of the new bowl involves the choice of the variant of the bowl shape or bowl decoration. For this reason, simulation iteration can be interpreted as corresponding to roughly one year, because this is close to the median use-life of serving pottery vessels, as the survey of the ethnoarchaeological literature shows (Mills 1989; Varien, Mills 1997). A simulation runs for 1000 iterations. Given the fact that archaeological site assemblages are not snapshots in time, but time-averaged, accumulated assemblages (Bailey 2007), the output of the simulation consists of matrices where sites are in rows and columns give the frequencies of variants accumulated in the last 200 iterations.

**Cultural scenarios**

In this paper, we present only two simple scenarios:

1. The Low Interaction scenario. All sites start with identical assemblages with maximum diversity (each entity has a different variant). Cultural transmission is based on the neutral model with the following parameters: $m = 0.1$, $\nu = 0.01$.

2. The High Interaction scenario. All sites start with identical assemblages with maximum diversity (each entity has a different variant). Cultural transmission is based on the neutral model with the following parameters: $m = 0.4$, $\nu = 0.01$.

**Methods of analysis**

**Comparing simulation results to the archaeological record**

Ideally, we would make a direct comparison of simulation results with the archaeological record by examining a correlation between the typological distance matrix produced by the simulation and the observed typological distance matrix based on archaeological data. However, the appropriate quantitative archaeological data is not available. This would also require simulated site positions to correspond to real site positions, which is not the case here.

For this reason, we attempt to make only an indirect and very rough comparison of simulated and real-world patterns using the traditional concept of archaeological culture as defined by Gordon Childe (1929) and formalised by David Clarke (1968). Childe defined cultures in this way: “We find certain types of remains – pots, implements, ornaments, burial rites and house forms – constantly recurring together. Such a complex of associated traits we shall call a ‘cultural group’ or just a ‘culture’. We assume that such a complex is the material expression of what today we would call ‘a people’” (Childe 1929.v–vi).

Clarke gave a formal version of this definition¹: “A polythetic set of specific and comprehensive artefact-type categories which consistently recur together in assemblages within a limited geographical area” (Clarke 1968.188).

The implication of Clarke’s formalisation is that archaeological cultures are equivalent to statistical groups (e.g., clusters resulting from cluster analysis). Culture historians have defined archaeological cultures as groups of sites sharing many types of material culture, with pottery usually being the most important class of material culture considered. The difference between cluster analysis and the traditional culture-historical approach is that definitions of traditional archaeological cultures are based on researchers’ subjective evaluations of similarities between assemblages. However, it is not unreasonable to assume that traditional cultures, despite the informal and subjective methodology used to define them, capture some of the main trends of spatial and temporal variability of material culture. It should be emphasised that our use of archaeological cultures as proxies for the patterns of material culture variability in space does not mean that we consider archaeological cultures as real anthropological and historical phenomena; we use them only as a proxy in the absence of quantitative data on the distribution of material culture.

So, in order to make our simulation results comparable to the distribution of archaeological cultures, we performed a cluster analysis on the assemblages using variant frequencies as variables. For frequency data, Euclidean distance matrix was calculated based on variant percentages in assemblages. Ward’s method was used as a clustering algorithm on variant percentage data (assemblages in rows, individual variant percentages in columns). The number of clusters is equated with the number of major archaeological cultures (4) in the study area: Vinča, Butmir, So-

---

¹ Childe was aware of the polythetic nature of archaeological cultures but his aim was to disregard types that were associated with more than one culture (Shennan 1994.13).
We have no simulated sites in the Adriatic area, so we did not anticipate the existence of clusters corresponding to the Danilo and Hvar-Lisići cultures. Given that both scenarios start with identical assemblages, the initial conditions are broadly comparable to the Early Neolithic in the Central Balkans, where only two major cultural groups are distinguished: Starčevo-Körös-Criş and Impresso group (Benac 1979). Since there are no sites in an area corresponding to the Impresso culture, the fact that all virtual sites have the same assemblage crudely mimics the relative uniformity of material culture in the Early Neolithic in the Central Balkans in the first half of the 6th millennium BC. The end of the simulation occurs 1000 iterations later, accumulating the variant frequencies from the last 200 iterations. The end of the simulation would roughly correspond to the first half of the 5th millennium BC, which is the period of the Late Neolithic in Central Balkans.

The results of cluster analyses were presented visually by plotting the sites coded for the cluster membership on the study area map. In this way, we can visually explore whether clusters based on typological similarity are spatially grouped in a similar way to traditionally defined archaeological cultures.

Is there a correlation between geographical and typological distances?

We answered this question by performing Mantel’s matrix correlation test (Mantel 1967) on the matrix of LCP distances and the matrix of typological distances based on type frequencies. We also tested for the correlation between typological distances based on the Jaccard distance measure and LCP geographical distances. The Jaccard distance measure...
is based on the Jaccard similarity coefficient (see below). The intensity and statistical significance of this correlation is also tested by Mantel’s matrix correlation test implemented in R.

Results

Patterns of typological similarity between simulated assemblages

Figure 4 shows the results for the Low Interaction scenario cluster analysis. There is a broad correspondence between the spatial patterning of simulated assemblages and the distribution of traditional archaeological cultures. Cluster 2 roughly corresponds to Vinča, cluster 4 to Butmir, Cluster 3 to Zelenikovo (although Zelenikovo is very similar to Vinča), and cluster 1 to Sopot-Lengyel, although it seems to have a discontinuous distribution – it is interrupted by cluster 2 sites in Vojvodina and its distribution then continues in Romania.

Figure 5 shows the plot of site locations with their scores on the first CA axis, which explains 1.9% of total variance (inertia) in the data. There is a gradient of CA axis 1 scores along the southeast-northwest axis. The similarity/dissimilarity pattern in space resembles the distribution of archaeological cultures. The group in Romania is clearly different from the group in Eastern Croatia. Bulgarian sites also form a clear group (shown in blue). The differences between Croatian and Bosnian sites are not as pronounced as the cluster analysis would suggest.

Figure 6 shows the results of the High Interaction scenario. In this plot, we recognise only a cluster of sites (Cluster 1) vaguely resembling Vinča culture, but the spatial distribution of other clusters is not that similar to the distribution of archaeological cultures.

Figure 7 shows the plot of site locations with their scores on the first CA axis, which explains 2.05% of total variance (inertia) in the data. Again, there is a gradient of CA axis 1 scores along the southeast-northwest axis.

Correlation of typological and geographical distances

Figure 8 shows the plot of LCP geographical distances and Euclidean typological distances, while Figure 9 shows the plot of LCP and Jaccard typological distances for the Low Interaction scenario; there is a curvilinear relationship between geographical and typological distances. Therefore, we transformed the data by calculating the logarithms of both variables. The correlation is higher for Jaccard distances (Pearson’s $r = 0.767$, Mantel test $p < 0.001$) than for Euclidean distances (Pearson’s $r = 0.265$, Mantel’s test $p < 0.001$).

Figure 10 shows the plot of LCP geographical distances and Euclidean typological distances, while Figure 11 shows the plot of LCP and Jaccard typological distances for the High Interaction scenario. Again, the correlation for log-transformed data is higher for Jaccard distances (Pearson’s $r = 0.839$, Mantel test $p < 0.001$) than for Euclidean distances (Pearson’s $r = 0.522$, Mantel test $p < 0.001$).

Discussion and conclusions

We have shown that a very simple model of interactions conditioned by distance can produce patterns similar to the spatial distribution of traditionally defined archaeological cultures. This does not mean that the neutral model is the correct model. It only

---

2 Even though this seems extremely small, it actually does summarize the patterns of covariation of main variants. There are around 20,000 variants in the accumulated assemblage, most of which are mutations with very low frequencies. Since they appear in very low frequency for a short time (usually being lost after the iterations in which they are introduced) they are mostly independent of the high frequency variants. Therefore, many of these variants are accounted by different CA axes thus reducing the percent of inertia explained by the first axis which accounts for a relatively small number of highly frequent variants.
means that this simple model is sufficient as it is to produce patterns that are visually similar to the observed ones, so we do not necessarily need to invoke ethnic identity or some other form of collective identity to account for the observed patterns. It is interesting that the gradient in typological similarity appears to be running in the NW-SE direction, although there is no directional movement or transmission.

The correlation between geographical and typological distances predictably differs between the High and the Low Interaction scenarios. The High Interaction scenario shows a stronger correlation than the Low Interaction scenario. However, in both scenarios a great amount of typological variance is not accounted for by distance, although the interaction is affected by distance. This means that we should not expect to find a perfect correlation between typological and geographical distances, even if the transmission is affected by distance (cf. Hart 2012), because the transmission process also occurs within sites.

It is interesting that the choice of the typological distance measure influences the strength of the correlation between geographical and typological distances. At present, we can only speculate about the reasons for this effect. Our first hypothesis was that this result could be explained by the fact that the Jaccard coefficient is predominantly influenced by joint presences of types. Given that joint presences are dependent primarily on the degree of interaction, which is conditioned by geographical distance, a typological distance measure based only on joint presences will capture mostly variation correlated with geographical distance. This is not the case for Euclidean distance measure which is based on all types: e.g., frequencies of types which are present only in one assemblage and not in the other, such as random innovations, will also enter into the calculation and act as random noise in relation to the signal created by interaction by geographical distance. However, we are not certain that this is the correct explanation, because when the ‘city-block’ distance metric is used, which is a distance analogue of the Brainerd-Robinson similarity coefficient (Brainerd 1951; Robinson 1951), the correlation between typological distances based on the city-block metric and geographical LCP distances for log transformed data are higher than for Euclidean typological distances, but not as high as for the Jaccard coefficient (e.g., for the Low Interactions scenario with city-block distance used: Pearson’s r = 0.556, p < 0.001). As Euclidean distance, the city-block distance also takes into account the frequencies of all types, regardless of whether they are jointly present or absent.

Fig. 6. Simulation results for the High Interaction scenario.

Fig. 7. Plot of simulated assemblages for the High Interaction scenario in geographical space coded for their score on the first CA axis.
While the results seem interesting, this project is still ‘under construction’, so there are many issues that need to be addressed and corrected before the simulation results can be considered relevant. One of the greatest weaknesses of the simulations described above is the lack of demographic dynamics. Demography is a key factor in the process of cultural evolution, as both theoretical work and empirical studies show (Collard et al. 2013; Henrich 2004; Shennan 2001; 2013). So, demographic dynamics needs to be built into the simulation framework – work that is currently in progress.

We simulated only a very simple and general (semi-abstract) model, so our results are interesting only as a methodological exercise at this point. The simulation approach is potentially useful because it allows for more complex and realistic models to be simulat-ed. For example, the model can be made more complex by letting the current similarity between sites affect transmission choices in addition to distance (cf. Axelrod 1997), or we could run simulations with several classes of ‘material culture’ coevolving and interacting. Additionally, spatial interaction can be modeled more flexibly (see Crema et al. 2014.292).

Needless to say, the full potential of simulations will not be fulfilled until empirical work has been done. Typological data from existing archaeological assemblages need to be collected in a systematic way in a form which is compatible with the simulation results. Only then will we be able to test different models and scenarios rigorously.
ACKNOWLEDGEMENTS

This research was undertaken as a part of project No. 177008, funded by the Ministry of Science and Technological Development of the Republic of Serbia. This paper is based on the paper presented at the 20th Neolithic seminar in Ljubljana. We are grateful to Professor Mihael Budja for providing us with the opportunity to participate in the seminar.

References


Simulating cultural transmission: preliminary results and their implications for the study of the formal variability ...


