Introduction to “Learning Strategies and Cultural Evolution During the Palaeolithic”

Kenichi Aoki and Alex Mesoudi

Abstract

In this introductory chapter, we first provide some background on the two major recurrent themes of the volume, i.e. learning strategies of individuals, and social and demographic characteristics of populations. This is followed by a brief summary of each chapter. Then, we conclude with some thoughts on why and how the methods and findings presented in this volume are relevant to, and might inform our understanding of, the replacement of Neanderthals by modern humans (Homo sapiens).

Keywords

Learning strategy • Demographic factors • Cultural change • Cultural diversity

This volume provides up-to-date coverage on the theory of cultural evolution as is being used by anthropologists, archaeologists, biologists, and psychologists to decipher hominin cultural change and cultural diversity during the Palaeolithic. The contributing authors are directly involved in this effort, and the material presented includes novel approaches and findings. The common theoretical framework of the volume is that cultural change constitutes a dynamic evolutionary system, which can be analyzed using tools and methods derived from the theory of biological evolution (Cavalli-Sforza and Feldman 1981; Boyd and Richerson 1985).

Various chapters show how learning strategies in combination with social and demographic factors (e.g. population size and mobility patterns) predict cultural evolution in a world without the printing press, radio, or the internet—which is to say that cultural traits can be acquired from others only by directly observing their actions or the results of these actions. Also addressed is the inverse problem of how learning strategies may be inferred from actual trajectories of cultural change, for example as seen in the North American Palaeolithic. Mathematics and statistics, a sometimes necessary part of theory, are explained in elementary terms where they appear, with details relegated to appendices. Full citations of the relevant literature will help the reader to further pursue any topic of interest.

1.1 Learning Strategies

Before proceeding it will be useful to briefly explain what the contributing authors and the editors mean by a “learning strategy.” A learning strategy is the way in which an organism combines individual learning and social learning, either simultaneously or sequentially, and its relative dependence on each. Here, individual learning occurs when the organism depends on personal experience to gather information, e.g. by trial-and-error. Social learning refers to obtaining information from other organisms, e.g. by imitation. Biases associated with social learning in the choice of whom to copy are also an integral part of a learning strategy.
Much theoretical work has been directed toward examining the adaptiveness of various social learning biases (Boyd and Richerson 1985; Laland 2004), such as “success bias” and “prestige bias” which entail preferentially copying a successful or a prestigious individual, respectively (Henrich and Gil-White 2001; Nakashashi et al. 2012), and “conformist bias” which entails copying the majority cultural behavior of one’s group (Henrich and Boyd 1998; Wakano and Aoki 2007; Nakashashi 2007; Kendal et al. 2009; see Aoki and Feldman 2014 for a comprehensive review). “Teaching” represents a powerful adjunct to social learning, where the individual being copied (the “teacher”) modifies his/her behavior to facilitate social learning by a naïve individual (the “pupil”) (Caro and Hauser 1992). Still another aspect of learning—in particular the learning of complex technical knowledge comprising various different skills—is that cultural traits may be interdependent, some serving as prerequisites for the acquisition of others. Several chapters in this volume tackle the question of how different learning strategies might structure population-level cultural change and variation, and the even more difficult problem of how to identify these population-level signatures in the often sketchy archaeological record.

1.2 Social and Demographic Factors

Many archaeologists and anthropologists currently emphasize social and demographic factors in interpreting “sudden” and “dramatic” changes in stone tools or other cultural artefacts during the Late Pleistocene (between 130,000 and 10,000 years ago), in particular the “creative explosions” (Kuhn 2012) of the African late Middle Stone Age and the European Upper Palaeolithic (Shennan 2001; Henrich 2004; Kline and Boyd 2010; Zilhão et al. 2010; Mesoudi 2011; Clark 2011; Kuhn 2013). In fact, theoretical studies have repeatedly shown that population size can have a large effect on cultural evolutionary rate and cultural diversity (Shennan 2001; Henrich 2004; Strumling et al. 2009; Mesoudi 2011; Lehmann et al. 2011; Aoki et al. 2011; Kobayashi and Aoki 2012; Aoki 2013), as can interconnectedness of subpopulations (Powell et al. 2009; Perreault and Brantingham 2011). Transmission chain experiments conducted in the laboratory also provide some support for a link between population (or group) size and cultural complexity (Derex et al. 2013; Muthukrishna et al. 2014; Kempe and Mesoudi 2014; but see Caldwell and Millen 2010).

However, archaeological evidence on the role of demographic factors is inconclusive or even contradictory. Two recent studies of Late Pleistocene South Africa are particularly relevant. Clark (2011) looked for signatures of population growth and/or demographic stress in an increase of diet breadth (e.g. the use of non-preferred prey animals), obtaining some support for an association with the heightened creativity of Howieson’s Poort. But, as Clark (2011) is careful to note, this association is open to an alternative interpretation, namely that rapid cultural change produced new tools, which were used to exploit novel resources. Klein and Steele (2013) (see also Klein 2008, Box 1) observed that edible shellfish remains from Middle Stone Age middens are significantly larger than those from Later Stone Age middens. If shellfish size reflects human collection intensity, then this finding suggests that the precocious appearance of modern behaviors in the Still Bay and Howieson’s Poort may not have been associated with population growth.

The claim that pre-contact Neanderthals in Spain used necklaces made of shells strung together as body ornamentation 50,000 years ago is also laden with ambiguity, in more ways than one. Zilhão et al. (2010) regards this as evidence for the cognitive equality of Neanderthals and modern humans, “support[ing] models of the emergence of behavioral modernity as caused by technological progress, demographic increase, …” However, perforation may not have been anthropogenic, and shells with naturally-formed holes of appropriate size for threading may have been selectively collected. Moreover, according to Prüfer et al. (2014), Neanderthal population size in the Altai region as estimated from genetic data shows a continual decrease after one million years ago, which is not true of various current modern humans. Similarly, Mellars and French (2011) argue for small population size in pre-contact European Neanderthals (MTA) compared to the Aurignacian. By implication, both Neanderthals and modern humans achieved the same cultural level, in spite of a difference in population size. Note, however, modern human beads occur much earlier—as early as 100,000 years ago in the Levant (Vanhaeren et al. 2006)—so perhaps population size did play a role.

In addition, statistical analyses of ethnographic hunter-gatherers have failed to detect an association between population size and the number of food-getting tools (Collard et al. 2005; Read 2006). On the other hand, ethnographic food-producing societies (e.g. small-scale farmers and herders) do conform to the theoretical prediction that population size and the number of food-getting tools should be positively correlated (Kline and Boyd 2010; Collard et al. 2013). Possible explanations for these contrasting results have been suggested, including higher degrees of specialization in the latter societies.

A fundamental problem in human evolution is how to account for an apparently abrupt cultural change, without invoking a major genetic change in cognition (e.g. innovativeness), for which there is at present no strong evidence (Klein 2008). Needless to say, absence of evidence does not constitute evidence of absence, and we are obliged to keep an open mind (Akazawa et al. 2013). Richerson et al. (2009)
(see also Richerson and Boyd 2013) discuss the possibility of spontaneous transitions between stable regimes—a small population at a low cultural level and a large population at a high cultural level. Developing this idea further and based on an explicit mathematical model, Aoki (submitted) shows that a salutary cultural change can be triggered by a gradual evolutionary change in the genetic basis for innovativeness. This scenario is not inconsistent with the “neural hypothesis,” a recent version of which invokes “a neural change that promoted the extraordinary modern human ability to innovate” (Klein 2008, p. 271). However, this neural change would not be attributable to just one “fortuitous mutation” in a major gene 50,000 years ago.

### 1.3 Summary of the Chapters

This volume comprises ten chapters, which use a range of methods to address different aspects of cultural evolution during the Palaeolithic.

In Chap. 2, Fogarty et al. present a theoretical analysis examining the modes and pathways of social learning, and how they affect the expected number of cultural traits maintained in a population. Specifically, they compare random oblique, best-of-$K$ (an example of direct bias, which entails a preference for a particular variant of a cultural trait), success bias, and one-to-many. Given the current emphasis among archaeologists and anthropologists on demographic factors, the effect of population size is also investigated, as is the less acknowledged role of innovation.

Fogarty et al. classify cultural traits as simple or complex, depending on the ease or difficulty of acquisition by social learning and innovation. Assuming an innate upper limit to the number of cultural traits that can be imagined—a limitation that may possibly be overcome by a mechanism analogous to “embedding” in linguistics—they show that the number of simple cultural traits may saturate as population size increases, in which case a statistical association between the two variables is not predicted. At smaller population sizes, there is a major effect of the mode of social learning. By contrast, the relation between the number of complex cultural traits and population size is approximately linear and almost identical for all modes of social learning investigated.

This is because most of the complex cultural traits that are maintained in the population can be accounted for by innovation alone, which raises the question of whether such traits qualify as “cultural” (Whiten et al. 1999).

In Chap. 3, Nakahashi describes and analyzes a new mathematical model for the evolution of teaching that is culturally transmitted rather than genetically determined. Teaching is here defined sensu Caro and Hauser (1992) as a knowledgeable individual (the teacher) altering its behavior in the presence of a naïve individual (the pupil), suffering a cost to do so, and thereby promoting social learning by that naïve individual. In this model, there are an infinite number of cultural traits, which are acquired by either individual learning or social learning, and where their acquisition entails a viability cost. Moreover, cultural traits are either beneficial or neutral, and only the former are assumed to contribute to fertility.

Nakahashi shows that teaching behavior can evolve culturally—i.e., teachers can invade and exist at a stable positive equilibrium—if a teacher can socially transmit more cultural traits than a non-teacher. However and surprisingly, it cannot evolve if teaching merely improves the accuracy of social learning by pupils. This latter result differs from the predictions of previous theoretical work that assume genetic determination of teaching behavior (Fogarty et al. 2011).

The next three chapters deal with structured populations. Kobayashi et al. (Chap. 4) directly address the cultural correlates of the replacement of Neanderthals (and other archaic humans) by modern humans. Their chapter begins with a detailed review of the archaeology of the Middle to Upper Palaeolithic transition in various parts of Eurasia, which suggests varying degrees of cultural continuity during/after the arrival of modern humans. In particular, China is apparently characterized by the late persistence of primitive core-and-flake industries (Norton and Jin 2009; Bar-Yosef and Wang 2012). Several Upper Palaeolithic industries in western Eurasia, e.g., the Emiran in the Levant and the Early Baradostian in the Zagros, may also exhibit recognizable elements of the preceding Middle Palaeolithic.

Kobayashi et al. describe a new model in which an invading modern human population has a demographic advantage (a higher relative growth rate), but receives unidirectional cultural influences from the indigenous archaic population. The cultural traits that the modern humans acquire from the archaics are assumed to be of a different kind from those that may be contributing to the demographic advantage of the former. Using approximate analytical methods and agent-based simulations, these authors show that biological replacement can be associated with either the rapid disappearance, the gradual disappearance, or the persistence of these autochthonous cultural traits. Gradual disappearance or persistence, i.e., cultural continuity, is predicted when a small modern human population invades a region with a relatively unfavorable physical environment. Importantly, cultural continuity is not an indicator of biological continuity.

The pattern of mobility within a geographically-structured population is recognized to be an important demographic factor in cultural evolution, through its effect on the variety of social learning opportunities (Powell et al. 2009). In addition, mobility may place a limit on the number of portable artefacts (Torrence 1983; Shott 1986). Premo (Chap. 5) gives an excellent introduction to residential mobility and logistical
mobility (Binford 1980; Kelly 1983), where the former refers to the relocation of a hunter-gatherer residential base and the latter to the movement of a subgroup on task-specific forays.

Premo conducts agent-based simulations of a spatially-distributed metapopulation to obtain its effective size. Each agent is a culturally monomorphic group of hunter-gatherers that relocates its residential base if a logistical foray proves unsuccessful in finding food. All agents initially carry different variants of a cultural trait, but agents within a certain interaction radius influence each other culturally, and the mean time until the metapopulation is fixed for one of these variants is used as a proxy measure of effective population size. The main result of this chapter is that the effective size of the metapopulation increases—sometimes quite substantially—as the frequency of residential moves decreases. This prediction has implications for the cultural diversity that is expected to be maintained in the metapopulation.

Madsen and Lipo (Chap. 6) describe and analyze a new agent-based simulation model for the cultural evolution of hierarchically-structured cultural traits (e.g. knowledge, skills), where some cultural traits are prerequisites for the acquisition of others. That is, while most models assume the transmission of independent cultural traits, Madsen and Lipo build on previous efforts (Mesoudi and O’Brien 2008; Mesoudi 2011; Enquist et al. 2011) to explore the more realistic situation where acquiring a cultural trait is dependent on already possessing other cultural traits. Cultural interactions, providing opportunities for social learning, can occur between two adjacent agents on a square lattice, as in the model proposed by Axelrod (1997). Agents may also innovate. The major difference between the current model and the original Axelrod model is that the focal agent can acquire a cultural trait from its neighbor only if it already has the prerequisite cultural traits. Alternatively, the neighbor may structure the learning environment of the focal agent by supplying the latter with the necessary prerequisite cultural traits. Madsen and Lipo refer to this behavior by the neighbor as teaching.

In their simulations, Madsen and Lipo examine the effects of the fidelity of teaching, size of design space (maximum possible number of cultural traits), innovation rate, and population size on cultural diversity (number of different repertoires of cultural traits segregating in the population), “knowledge depth” of cultural traits (average number of prerequisites per extant cultural trait, which can be regarded as a measure of cultural complexity or cumulativeness), etc. Cultural diversity is found to increase with the fidelity of teaching and with the size of design space, but only when innovation is allowed. Knowledge depth increases with the fidelity of teaching, but again only when innovation is allowed.

Of particular interest is the possibly counterintuitive prediction that knowledge depth is minimally affected by population size. This result is likely a consequence of the assumptions inherited from Axelrod (1997) that cultural interactions are spatially localized and moreover occur preferentially between culturally similar agents (homophily). If these assumptions apply to Palaeolithic hunter-gatherers—they may not to some present-day hunter-gatherers that travel long distances (Hewlett et al. 1982)—then the predictive value of population size per se should be viewed with caution.

Chapters 7 and 8 explore novel Bayesian methods for detecting signatures of learning strategies in archaeological data. Summary statistics such as the expected number of cultural traits are useful, but frequency distributions (spectrums) of variants of cultural traits are more informative. Kandler and Powell (Chap. 7) explain a powerful new method—approximately Bayesian computation (ABC) (Beaumont et al. 2002)—for identifying learning strategies that produce the observed population level data on such frequency distributions. Very briefly, ABC entails simulating a model with parameter values chosen from a prior distribution and retaining those parameter values that give the closest fit of the simulated data to the observed data; these retained parameter values approximate the posterior distribution.

As a concrete example of the application of this method, they first generate “observed” data by simulating a hypothetical model for the evolution of a cultural trait with known parameter values but with noise added. The parameters that define the learning strategy are the degrees of reliance on individual learning (i.e. innovation), directly-biased social learning, and conformist social learning. Then, the frequencies of the variants are sampled at various times, and the parameters of this model are estimated by ABC. It is shown that the original parameter settings are faithfully recovered by this estimation procedure.

Kovacevic et al. (Chap. 8) apply approximate Bayesian computation to empirical data on the geographical distribution of bead types in European Aurignacian sites. The goal of this chapter is to test the hypothesis, proposed by Vanhaeren and d’Errico (2006) based on this data, that these bead types had a symbolic meaning and served as markers of ethnic identity. Agents in the simulation model of Kovacevic et al. are mobile Aurignacian groups, which can undergo pairwise cultural interactions when in geographical proximity. Two cultural interaction processes are considered, “conflict” and “sharing.” In the case of conflict, the bead types of the losing group are completely replaced by those of the winning group. Sharing entails the pooling and swapping of some bead types between the two groups.

If bead types are indeed ethnic markers, then—as Kovacevic et al. argue—two interacting groups that are relatively similar for bead types are more likely to share, whereas those that are relatively different are more likely to experience
conflict (culture-dependent interaction model). On the other hand, if bead types have no such meaning, then the nature of the cultural interaction should not depend on these similarities or differences (null model). Kovacevic et al. find that the best fits of the simulated to the observed data are equally likely under the null model as the culture-dependent model. Thus, their study does not support the hypothesis that the Aurignacian was ethnically structured, at least in the sense that different bead types were used symbolically to mark ethnic identity.

In Chap. 9, O’Brien et al. provide an excellent summary of the variety of learning strategies discussed in the literature and of the well-studied archaeology of Palaeolithic North America. Among these various learning strategies, they focus on “guided variation” and “indirectly-biased” social learning as likely candidates for the Early Paleoindian period. Guided variation is a learning strategy in which unbiased social learning is followed by individual learning that targets the environmentally optimal behavior (Boyd and Richerson 1985). Indirect bias entails that an individual perceived to be successful or prestigious is preferentially copied. Whereas the former results in adaptive cultural change, the latter may not do so unless success or prestige is correlated with biological fitness.

O’Brien et al. review several recent studies (Morrow and Morrow 1999; Buchanan and Hamilton 2009; Hamilton and Buchanan 2009; Shohts et al. 2012; Buchanan et al. 2014) that ask whether the geographical variation in Clovis points is due to regional adaptation by guided variation or other factors such as random drift associated with indirectly-biased social learning. They argue that different learning strategies may be applied to different aspects of stone-tool production, specifically that patterns of flake removal may have been determined by prestige-biased social learning from skilled craftsmen, whereas point shape was determined by guided variation. The greater regional variation observed for point shape is explained by this difference in learning strategies.

Our final two chapters focus on laboratory experiments as a means of better understanding learning strategies, beginning with Caldwell (Chap. 10). The distinguishing feature of the culture of modern humans as opposed to non-human animal cultures—to a certain extent, perhaps also Neanderthal culture—is that it is cumulative. Thus, as Tomasello (1999, p. 512) explains, “[t]he most distinctive characteristic of human cultural evolution is the way that modifications to a cultural artifact or a social practice made by one individual or group of individuals often spread within the group, and then stay in place until some future individual or individuals make further modifications.”

Caldwell reviews her work on transmission-chain laboratory experiments (e.g. Caldwell and Millen 2008a, b, 2009, 2010; Caldwell et al. 2012), which ask what unique aspects of cognition and social learning in present-day humans make cumulative cultural change possible. One prevailing view holds that only imitation, i.e. “process-oriented” or “action-copying” social learning, can support high-fidelity social learning, which is a necessary condition for cumulative cultural change. Against this, the experiments tested whether emulation, i.e. “goal-oriented” or “results-feedback” social learning, might also qualify.

Participants were given two tasks, building a paper airplane scored for flight distance and building a spaghetti tower scored for height. The experimental conditions differed in the type of information made available to the participants, specifically the actions used by the antecedents in the transmission chain to execute the tasks and/or the results (i.e. products) of these actions. Caldwell and coworkers found that high-fidelity social learning occurred even when participants were permitted to observe only the results. Moreover and consequentially, performance was improved over the “generations” of the transmission chain. Teaching was also found to be effective, independently of imitation and emulation. However, as Caldwell notes, these experiments do not address the role of innovation, which is clearly also a necessary condition for cumulative cultural change (Borenstein et al. 2008).

Furthermore, the tasks used are rather simple compared to even very early hominin technology such as flint-knapped handaxes, for which imitation may well be more important.

Lycett et al. (Chap. 11) review their work on transmission-chain experiments conducted in the laboratory (Kempe et al. 2012; Schillinger et al. 2014). They emphasize the importance of laboratory experiments in understanding how the dynamics of micro-evolutionary processes affect artefactual variation, to produce the macro-patterns seen in the archaeological record. Specifically, their interest focuses on elucidating the relation between copying error, either deliberate or unintentional, and cultural variation among populations.

In the first experiment (Kempe et al. 2012), each participant was asked to view an image of a handaxe drawn by his/her antecedent in the transmission chain and to faithfully copy its size. Observed copying error averaged 3.43 %, which is consistent with the value of 3 % reported in the psychophysical literature. Moreover, the variance among transmission chains increased as expected over the ten generations of the experiment. Then, individual based simulations were conducted to predict the variance of handaxe length and breadth after 200 generations. The predicted variance was far greater than that observed in a sample of more than 2,000 Acheulean handaxes from 21 sites spanning 1.2 million years, suggesting certain undetermined factors countermanding the effects of copying error.

In the second experiment (Schillinger et al. 2014), participants were asked to make a replica handaxe from plasticine using a knife. Two experimental conditions were compared: reductive only in which material may be removed but not added as is usually the case in stone-tool knapping, and...
additive-reductive in which both actions are permitted as in the manufacture of pottery. As expected, copying error was statistically greater under the former condition. These simple experiments show, using a “model artifact” for culture—in this case, handaxes—equivalent to “model organisms” in experimental population genetics, how features of perceptual systems and manufacturing techniques can generate systematic patterns in cultural datasets.

1.4 Concluding Remarks

The goal of this volume is to show how learning strategies, in combination with social and demographic factors, predict cultural change and cultural diversity during the Palaeolithic. Also addressed is the more challenging inverse problem of how learning strategies may be inferred from the sketchy archaeological record. Towards this end, our contributors have applied a diverse range of scientific methodologies, including mathematical models derived from population genetics, spatially explicit agent-based models, approximate Bayesian computation, and transmission chain laboratory experiments.

Our underlying motivation is to explain why and how Neanderthals were replaced (or assimilated) by modern humans, in terms of cultural differences between the two (sub-)species. This is an entirely reasonable premise, given that replacements in historical times were likely driven by cultural differences—and the demographic or disease-related corollaries of such cultural differences—between competing ethnic groups (Diamond 1997). Moreover, as recently as 2003, leading anthropologists noted that “[t]he ultimate mechanism for this replacement is widely considered to be a behavioral difference between non-modern and modern populations that lent an adaptive advantage to moderns” (Henshilwood and Marean 2003, p. 627). The theory described in this volume helps us to understand how various factors—inventiveness, biases and error associated with social learning, population size and structure, residential and logistical mobility, to name a few—influence the amount, complexity, and geographic variation of culture.

However, based on intensive archaeological work during the past decade, it has been suggested that contemporaneous Neanderthals and modern humans—in particular in Europe where the two (sub-)species apparently overlapped for several thousand years (Higham et al. 2014)—both exhibited most, perhaps all, “modern behaviors” including symbolic behavior (d’Errico and Stringer 2011; Zilhão 2013). See, for example, Table 3 of McBrearty and Brooks (2000) for a comprehensive list of modern behaviors, initially believed to constitute the distinguishing features of modern humans. From the standpoint of the cognitive equality of Neanderthals and modern humans, the possible “nail in the coffin” is the recent report of a rock engraving made by pre-contact Neanderthals in Gibraltar (Rodríguez-Vidal et al. 2014).

On the other hand, we must not forget that there was much variability both among Neanderthal regional groups and among modern human regional groups. Competition resulting in replacement would likely have occurred on a circumscribed geographical scale, not between the “champions” on either side. Hence, if Neanderthal regional cultures were on average slightly inferior to modern human regional cultures—perhaps reflecting a small (sub-)specific difference in cognition or in demography—then the premise that cultural differences contributed to replacement, which we subscribe to, would still be tenable.

References


Aoki K (submitted) Modeling abrupt cultural regime shifts during the Palaeolithic/Stone Age


Introduction to “Learning Strategies and Cultural Evolution During the Palaeolithic”


