

THE CULTURAL TRANSMISSION OF GREAT BASIN PROJECTILE-POINT TECHNOLOGY I: AN EXPERIMENTAL SIMULATION

Alex Mesoudi and Michael J. O'Brien

A Darwinian evolutionary approach to archaeology naturally leads to a focus on cultural transmission. Theoretical models of cultural evolution indicate that individual-level details of cultural transmission can have specific and significant population-level effects, implying that differences in transmission may be detectable in the archaeological record. Here we present an experimental simulation of the cultural transmission of prehistoric projectile-point technology, simulating the two transmission modes—indirect bias and guided variation—that Bettinger and Eerkens (1999) suggested were responsible for differences in Nevada and California point-attribute correlations. Groups of participants designed “virtual projectile points” and tested them in “virtual hunting environments,” with different phases of learning simulating, alternately, indirectly biased cultural transmission and independent individual learning. As predicted, periods of cultural transmission were associated with significantly stronger attribute correlations than were periods of individual learning. We also found that participants who could engage in indirectly biased horizontal cultural transmission outperformed individual-learning controls, especially when individual learning was costly and the selective environment was multimodal. The study demonstrates that experimental simulations of cultural transmission, used alongside archaeological data, mathematical models and computer simulations, constitute a useful tool for studying cultural change.

Una perspectiva Darwiniana de la arqueología conduce, naturalmente, hacia un enfoque en la transmisión cultural. Los modelos teóricos de la evolución cultural indican que el nivel individual de la transmisión de cultura puede tener efectos específicos y significativos sobre el nivel demográfico. Esto implica que las diferencias de transmisión pueden ser detectadas en la documentación arqueológica. Aquí presentamos una simulación experimental sobre la transmisión cultural de la tecnología prehistórica de puntas de proyectil, representando los dos modos de transmisión: tendencia indirecta y variación guiada. Ambos modos, según Bettinger y Eerkens, fueron responsables de las diferencias en las correlaciones punto-parámetro de Nevada y California. Grupos de participantes diseñaron “puntas virtuales de proyectil” y las probaron en “medios virtuales de cacería” con diferentes fases de aprendizaje, simulando el modo de tendencia indirecta y el aprendizaje individual independiente. Como fue previsto, los periodos de transmisión cultural fueron asociados con parámetros de correlaciones significativamente más fuertes que los periodos de aprendizaje individual. También encontramos que los participantes dedicados a la transmisión cultural horizontal de tendencia indirecta, superaron los controles de aprendizaje individual. Este resultado se obtuvo especialmente entre los grandes grupos, cuyo aprendizaje individual es costoso y el entorno selecto es multi-modal. El presente trabajo demuestra que las simulaciones experimentales de transmisión cultural, usadas con modelos matemáticos y simulaciones en computadoras constituyen una herramienta para estudiar los cambios culturales.

There has been growing interest in recent years in applying the theoretical and methodological tenets of Darwinian evolutionary theory to the archaeological record (e.g., Barton and Clark 1997; Dunnell 1980; Lipo et al. 1997, 2005; Lyman and O'Brien 1998; Neiman 1995; O'Brien 1996; O'Brien and Lyman 2000, 2002, 2003a; Shennan 2002; Shennan and Wilkinson

2001). One of the key theoretical tenets of any theory of Darwinian evolution, be it biological or cultural, is *inheritance*: “Any variation which is not inherited is unimportant for us” (Darwin 1859:75). In cultural evolution (Boyd and Richerson 1985; Cavalli-Sforza and Feldman 1981; Mesoudi et al. 2004, 2006; Richerson and Boyd 2005), inheritance takes the form of *cultural transmission*, the process

Alex Mesoudi ■ Department of Anthropology, University of Missouri, Columbia, MO 65211. Current Address: Department of Social and Developmental Psychology, University of Cambridge, Free School Lane, Cambridge CB2 3RQ, United Kingdom (am786@cam.ac.uk)

Michael J. O'Brien ■ Department of Anthropology, University of Missouri, Columbia, MO 65211 (ObrienM@missouri.edu)

by which information (e.g., knowledge, skills, or beliefs) is passed from individual to individual via social learning. Consequently, evolutionary archaeologists have begun to seek signs of cultural transmission in the archaeological record (O'Brien and Lyman 2003b). O'Brien and Lyman (2000) and Lipo et al. (1997) have argued that seriations constitute temporal sequences of artifacts because those artifacts are causally linked by cultural transmission (they exhibit "heritable continuity"), while they and others have applied phylogenetic analyses to identify artifacts that are related by cultural transmission (descent) or that represent independent innovations in unrelated lineages (Collard et al. 2006; Lipo et al. 2005; O'Brien et al. 2001; O'Brien and Lyman 2003a).

Although a Darwinian evolutionary approach to archaeology naturally leads to a focus on transmission, it does not necessarily follow that cultural transmission must be identical in its details to biological inheritance; in many cases cultural transmission will exhibit its own unique rules and dynamics. These differences do not invalidate an evolutionary approach to archaeology—all that is required for Darwinian evolution is that transmission, however it is realized, takes place. However, there is one general rule that archaeologists would benefit from taking from evolutionary biology—that the small-scale details of inheritance and selection at the individual level may have dramatic effects at the population level. In evolutionary biology, this link between microevolution and macroevolution was made during the 1930s and 1940s, when the microevolutionary principles of theoretical and experimental population genetics were used to explain temporal and spatial macroevolutionary patterns found by paleontologists and naturalists. Indeed, it was not until this "Modern Synthesis" (Huxley 1942; Mayr and Provine 1980) occurred that evolutionary biology became the hugely successful discipline that it now is. The same general principle—that microevolutionary processes generate macroevolutionary phenomena—also applies to cultural evolution.¹ Models of cultural evolution (e.g., Boyd and Richerson 1985, 2005; Cavalli-Sforza and Feldman 1981) have highlighted how the details of cultural transmission at the individual level frequently have important effects at the population level. For example, Boyd and Richerson (1985) proposed a num-

ber of biases in cultural transmission, such as "guided variation," in which individuals acquire a cultural trait, then modify it through individual trial and error; "conformist bias," in which the most popular variant in a population is disproportionately more likely to be adopted; and "indirect bias," in which traits exhibited by successful or high-status individuals are preferentially adopted. These biases each generate specific population-level effects (Boyd and Richerson 1985), such as widespread cooperation generated by cultural group selection (conformist bias) or runaway prestige markers (indirect bias). This theoretical work has important implications for archaeology. Given that archaeologists predominantly study large samples of population-level data that were generated over extended periods of time (i.e., cultural macroevolution: Mesoudi et al. 2006), it follows that this population-level data will, in many cases, be affected by and structured according to individual-level details of cultural transmission (i.e., cultural microevolution: Mesoudi et al. 2006). By making explicit assumptions and predictions regarding cultural transmission, archaeologists can gain better insight into past cultural change, and better explain specific patterns and trends in the archaeological record.

Some archaeologists have already taken advantage of cultural transmission theory to help explain certain patterns in the archaeological record. One example is Bettinger and Eerkens's (1997, 1999) study of Great Basin projectile points manufactured around A.D. 300–600 following the replacement of the atlatl with the bow and arrow. Bettinger and Eerkens (1999) observed that points found in two regions of the Great Basin differ in the degree to which their attributes, such as weight, width, and length, correlate with each other (Table 1). They attribute these differences to the manner in which prehistoric people of the two regions acquired and transmitted projectile-point technology. Specifically, the attributes of points found in eastern California were found to be poorly correlated with each other, which Bettinger and Eerkens (1999) argued was because point designs in this region originally spread as a result of guided variation. Hence, each attribute was subject to separate individual trial-and-error experimentation, causing them to vary independently. In contrast, projectile points of the same material and from around the

Table 1. Correlation Coefficients for Great Basin Point Attributes, by Region (Nevada or California)

		Axial length	Max. width	Basal width	Neck width	Thickness	Distal shoulder angle	Proximal shoulder angle	Weight
Max. length	NV	1.00	.58*	.48*	.51	.51*	-.19	-.34	.86
	CA	1.00	.24	.07	.21	.15	-.24	-.24	.76
Axial length	NV		.59	.49*	.51*	.49*	-.18	-.33	.86
	CA		.22	.04	.16	.14	-.24	-.27	.74
	Max. width								
			NV	.44	.72	.12	-.48	-.53*	.62
			CA	.54	.69	.26	-.47	-.08	.71
			Basal width						
				NV	.71	.21	.11	.26	.80*
				CA	.78	.05	-.06	.31	.14
				Neck width					
					NV	.23	-.15	-.15	.61
					CA	.11	-.06	.15	.40
					Thickness				
						NV	.14	-.02	.70
						CA	.03	-.02	.57
						Distal shoulder			
							NV	.53*	-.12
							CA	.16	-.38
							Proximal shoulder		
								NV	-.21
								CA	-.01

Note: Correlations (Pearson's r) from Bettinger and Eerkens (1999:237).

NV = Nevada sample, CA = California sample.

*Significantly stronger correlation in the Nevada sample than in the California sample.

same period found in central Nevada featured uniform designs with highly correlated attributes. Bettinger and Eerkens (1999) argued that points in this region originally spread as a result of indirect bias, with individuals copying wholesale the design of a single successful model.² Hence, differences at the individual level (guided variation vs. indirect bias) can be argued to have generated differences at the population level (uncorrelated attributes vs. correlated attributes).

One limitation inherent within archaeology is that we usually have access only to population-level historical data. The details of cultural transmission at the level of the individual—who copies what from whom, and how—can in most cases only be inferred from these archaeological data, as Bettinger and Eerkens (1999) did, and seldom directly observed or measured. Mathematical simulations offer one means of addressing this problem, with the results of simple models of cultural transmission matched to archaeological data (e.g., Eerkens et al. 2005). Mathematical models, however, are only as good as their assumptions, in this case assumptions regarding people's propensities to learn socially rather than individually, to conform, to copy the most successful individual, and so on. We also need experimental data from psychology in order to verify the assumptions and findings of theoretical models. Although Boyd and

Richerson (1985) attempted to ground their models in empirical data from psychology, sociology, and anthropology, as they themselves acknowledged (Boyd and Richerson 1985:296), many of these data were inadequate and/or not collected with their specific hypotheses in mind. Indeed, the mismatch between the recent findings of experimental economics (e.g., Henrich et al. 2004) and the assumptions of traditional theoretical economics illustrates the dangers of failing to test theoretical models with experimental data. Experimental methods also allow the manipulation of variables, access to complete, uninterrupted data, and the comparison of fitness at different points in time, all of which are impractical or impossible with archaeological methods, making experiments a potentially valuable tool for understanding archaeological phenomena (Mesoudi 2008). Indeed, experimental laboratory studies have made major contributions to evolutionary biology by uncovering the fundamental principles of biological inheritance and selection, from Mendel's (1866) pea-plant studies, which established that inheritance is particulate, to Morgan's (1932) and Dobzhansky's (1937) groundbreaking experiments of selection and mutation in *Drosophila*, through to present-day experimental population genetics. Recent studies have experimentally simulated long-term evolution using microorganisms (Elena and

Lenski 2003; Lenski and Travisano 1994), tackling issues regarding macroevolution that previously only the fossil record and/or theoretical models could address, such as punctuated equilibria (Lenski and Travisano 1994), long-term adaptation in multimodal fitness landscapes (Colegrave and Buckling 2005; Elena and Lenski 2003), and the evolution of sexual reproduction (Colegrave 2002). Psychological experiments that simulate different modes of cultural transmission have the potential to be similarly useful in explicating past cultural evolution (Mesoudi 2007, 2008). Although experimental studies are common in archaeology, often specifically regarding projectile points and related projectile technology (e.g., Cheshier and Kelly 2006; Flenniken and Raymond 1986; Knecht 1997; Odell and Cowan 1986; Schiffer and Skibo 1987), these studies deal exclusively with the physical properties of different projectile technologies by simulating their use under different conditions. To our knowledge, no experimental studies have attempted to simulate the cultural transmission of prehistoric artifacts, which the models of Boyd and Richerson (1985) and others, and the analyses of Bettinger and Eerkens (1997, 1999), suggest played an important role in generating systematic patterns in the archaeological record.

Although not specifically addressing archaeological issues, there have recently been attempts to experimentally test the assumptions and findings of the mathematical models of cultural transmission mentioned above. McElreath et al. (2005) had groups of 6–10 participants repeatedly play a computer-based task (planting one of two crops, one of which gave a higher yield), either with no social learning (an individual-learning baseline), or with the opportunity to view the previous choice of one randomly selected group member (allowing simple social learning), or with the opportunity to view the previous choices of all group members (allowing conformity). Although some participants did engage in social learning, most did not, even when mathematical models predicted that social learning would have been optimal. Of those who did, social information was more likely to be used when individual learning was relatively inaccurate and the environment (which crop was optimal) did not change, in line with theoretical predictions. Although models indicated that conformity was the optimal strategy under all conditions, it was

used only when the environment fluctuated. In another experiment, Kameda and Nakanishi (2002) explored the adaptiveness of individual and social learning within groups of six participants repeatedly playing a simple computer-based task (locating a rabbit in one of two possible locations). Participants had the opportunity to engage in either individual learning at a cost or free social learning, in which the past choices of three randomly selected group members could be seen. In accordance with the predictions of the authors' mathematical simulations, polymorphic equilibria emerged in which individual learners ("information producers") and social learners ("information scroungers") coexisted. Increasing the cost of individual learning increased the frequency of social learning. A follow-up study (Kameda and Nakanishi 2003) found that groups in which social learning was possible outperformed groups in which only individual learning was allowed, suggesting that social learning is adaptive when individuals can selectively switch between individual and social learning.

The aim of our study was to test Bettinger and Eerkens's (1999) hypothesis that the different patterns of projectile-point variation observed in the Great Basin are the result of different cultural transmission processes—guided variation and indirect bias—by experimentally simulating those cultural-transmission processes in the laboratory using similar experimental methods to those just outlined. Participants played a simple computer game in which they tested "virtual projectile points" in a series of "virtual hunting environments." Learning was divided into separate phases (Table 2). In Phase 1, participants were given a list of players from a previous group together with each previous player's overall success rate, and they were asked to copy the point design of one of those players. Here we predict that participants will predominantly choose to copy the most successful model (exhibit indirect bias), causing the attributes (e.g., length and width) of the points to correlate, as in Bettinger and Eerkens's (1999) central-Nevada data. In Phase 2, participants were allowed to experiment with their points by changing their attributes, with no access to any other player's design (simulating guided variation). Here we predict that the correlations between the attributes will decrease, as in Bettinger and Eerkens's (1999) eastern-California data. If the

Table 2. The Three Learning Phases and Associated Predictions.

Phase 1	Phase 2	Phase 3
Hunt 1	Hunts 2-25	Hunts 26-30
Oblique Cultural Transmission	Individual Learning	Horizontal Cultural Transmission
Description: Participants saw a list of six previous participants along with those participants' final scores. Participants selected one of the previous participants and copied their point design. No modification was allowed.	Description: Participants could modify any of the five point attributes. Participants had no access to anyone else's point design.	Description: Participants could modify their point designs as in Phase 2. Participants could additionally see the cumulative scores of every other participant in their group, and could view the point design of one of those other participants.
Prediction: Strong Attribute Correlations	Prediction: Weak Attribute Correlations	Prediction: Strong Attribute Correlations
Prediction: High Within-Group CV	Prediction: Low Within-Group CV	Prediction: High Within-Group CV

experimental data generated by our participants match up with the archaeological data of Bettinger and Eerkens (1999), then we can be more confident in their claims. If the historical and experimental data sets do *not* match, then identifying exactly why they do not match would nevertheless be a productive exercise, either in calling Bettinger and Eerkens's (1999) hypothesis or methods into question or in identifying exactly what is missing from the experimental simulation.

Bettinger and Eerkens (1999) do not specify why point designs would diverge during the period of individual learning. It may have been that designs were stylistic and diverged through a process of drift. Here we test an alternative hypothesis that point designs were functional—that the values of attributes such as length or width contributed to a point's efficacy in hunting. This is consistent with experimental studies such as that of Cheshier and Kelly (2006), who found that point attributes (e.g., the thickness:length ratio) affect performance characteristics of the point when fired. If point designs are functional, then there would be no reason to modify a point inherited from a successful model unless the selective environment has changed. Hence, in Phase 1 the participants experienced different selective "environments" than those of the models. The point inherited from the most successful model in the previous group was therefore not necessarily the best point in the participant's environment, providing an incentive for the participant to experiment with the design. There were three different environments (E1, E2, and E3), all different from the environment of the models (E0), so we predict that during Phase 2 points will diverge, and correlations between attributes will decrease, as a result of adapting to these different environments. A second possible reason for diver-

gence during individual trial and error is that points in the same environment may have had different locally optimal designs as a result of tradeoffs between competing uses or requirements. Hence, three continuous attributes (Length, Width, and Thickness) had bimodal fitness functions such as the ones shown in Figure 1, featuring a global optimum (giving the maximum possible score from that attribute) and a local optimum (giving two-thirds the payoff of the global optimum). Any minor deviation away from either of these two optima reduced the participant's overall score, which was the sum of each attribute's fitness. Combining three such bimodal fitness functions creates a multimodal "adaptive landscape." This is a concept commonly used in evolutionary biology (Arnold et al. 2001; Simpson 1944; Wright 1932) to represent the design space of all possible combinations of multiple phenotypic characters, where the height of the landscape represents fitness. Similarly, our adaptive landscape represents all possible point designs, or all possible combinations of all values of Length, Width, and Thickness, where the height of the landscape represents the payoff to the participant of that point design. So, for example, a point specified by values of Length, Width, and Thickness that are all at global optima would be found at the highest peak in the landscape; a point specified by one global and two local optima would be at a slightly lower peak; and a point at all three local optima would be found at the lowest peak (there are eight peaks in total). Hence, diversity will be produced and maintained during Phase 2 when participants converge on different locally optimal peaks. Two other attributes, Shape and Color, were discrete. Shape had a linear fitness function (Figure 1), whereas Color was neutral and did not affect feedback.

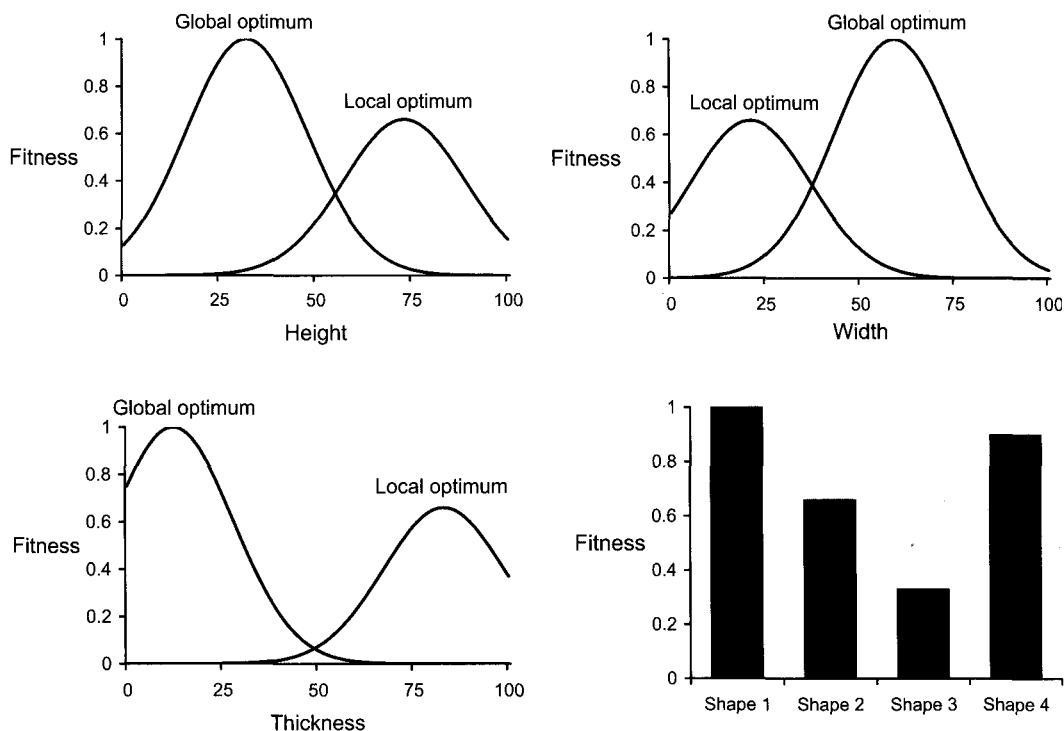


Figure 1. Example fitness distributions for the functional point attributes.

The assumption of a multimodal adaptive landscape underlying point design allows us to test a more general hypothesis regarding cultural transmission—that indirectly biased cultural transmission allows individuals to jump from locally optimal peaks in the design space to the globally optimal peak (or higher locally optimal peaks). Participants relying solely on individual learning will explore the design space by improving their scores (moving uphill) but are likely to end at a globally suboptimal peak. This is a common finding in experimental studies of individual learning and problem solving (e.g., Fu and Gray 2004, 2006). In a sufficiently large group, however, one or more participants might by chance reach the globally optimal peak and outperform the other group members. If the participants are able to engage in indirectly biased cultural transmission, then they can potentially jump from their locally optimal peak to the globally optimal peak (or a higher locally optimal peak) of the most successful individual.³ In order to test this hypothesis, in Phase 3 participants were allowed to copy the designs of their fellow group members (given information about those

members' success in the task). This potentially allows participants to jump from the locally optimal peak that they found during Phase 2 to the globally optimal peak (or the highest peak found by the participants in that group). We therefore predict that participants in groups will outperform individual controls during this phase of learning. We also predict that this indirectly biased horizontal cultural transmission in Phase 3 will have the same effect as indirectly biased oblique cultural transmission in Phase 1, i.e., increase the correlations between point attributes. Indeed, this might constitute a stronger test of Bettinger and Eerkens's hypothesis because in Phase 3, unlike in Phase 1, participants were not forced to copy the point design of a model.

Several other issues were also investigated. In the third of three seasons of hunting, participants had to pay a cost to modify their point designs. Costly individual learning is predicted to increase the frequency and adaptiveness of social learning (Boyd and Richerson 1989, 1995), consistent with previous experiments (Kameda and Nakanishi 2002; McElreath et al. 2005). This may, therefore,

provide an explanation for why prehistoric hunters in Nevada and California differed in their reliance on social learning. We are also interested in comparing variation in point designs between and within different environments before and after horizontal cultural transmission, predicting that indirectly biased within-group horizontal cultural transmission will reduce point variation within groups while not affecting between-group variation. That is, if every member of a group copies the single most-successful individual in their group, then everyone in that group will soon have similar point designs, resulting in low within-group variation. If cultural transmission is only permitted within groups (i.e., people cannot copy individuals from other groups), and the most successful individuals in each group have different point designs (due to the different environments), then each group will converge on a different design, thereby maintaining between-group variation. This constitutes a test of Eerkens and Lipo's (2005) recent suggestion that changes in variation can be used to identify different forms of cultural transmission in archaeological data,⁴ as well as potentially provide the conditions for cultural group selection (Boyd and Richerson 1985; Richerson and Boyd 2005).

Experimental Methods

Participants

Eighty-one students at the University of Missouri–Columbia participated in the study (54.3 percent female; mean age = 23.15, *sd* = 7.08). Each participant was paid \$8 for completing the approximately 45-minute experiment. Participants were recruited through advertisements in the student newspaper, fliers posted around campus, and postings on a University of Missouri email bulletin board. Participants were unaware of archaeological work regarding projectile points in order to prevent prior knowledge interfering with point design. Six participants completed the pretest phase. The remaining participants (*n* = 75) took part either on their own (*n* = 24) or in groups of two (*n* = 6), three (*n* = 9), four (*n* = 8), five (*n* = 10), or six (*n* = 18). Participants/groups were randomly assigned to Environment 1 (*n* = 27), Environment 2 (*n* = 19), or Environment 3 (*n* = 29).

Task and Design

The experiment consisted of an interactive computer game programmed in Borland C++ Builder 6.0 and implemented on a local network of six desktop PCs. Participants designed a series of "virtual projectile points." Each participant completed three seasons of hunting, with each season consisting of 30 trials, or "hunts." One hunt represents one opportunity to modify the point design and to test it in the virtual hunting environment. A projectile-point design was represented by five values, each constituting a different attribute. Length, Width, and Thickness were continuous variables of arbitrary units ranging from 1 to 100. Shape and Color were discrete variables, each taking one of four values ("Shape 1," "Shape 2," "Shape 3," or "Shape 4," and "White," "Gray," "Fawn" or "Orange," respectively). Generic shape terms were used because participants were not familiar with technical terms such as "corner notched" or "lanceolate." The values of these five attributes were used to calculate the "fitness" of the design, in terms of the number of calories (from 1–1000) it yielded during a hunt (see below for fitness equations). "Fitness" here is intended as a loose term indicating the functional efficacy of a particular design, although given that projectile technology was originally used to obtain food and can be seen as part of an individual's "extended phenotype" (O'Brien and Holland 1995), fitness may also be interpreted in a biological sense of survival and reproduction.

Before the experimental sessions began, six participants completed a pretest study during which they completed one season of 30 hunts using purely individual learning, with no access to the design of any other participant. These six pretest participants served as "models" for the first phase of the experimental sessions. During the experimental sessions, every participant played three hunting seasons each comprising 30 hunts, with each season of 30 hunts divided into three phases (Table 2).⁵ During the first hunt (Phase 1), which represents oblique cultural transmission, participants could see on the screen a list of the six pretest participants (the models) and each model's final score, cumulative over the 30 hunts. The participant selected one of these models to copy, after which the model's final point design was copied to the experimental participant's screen. No modification of the point design was permitted at this stage. After a model was selected, par-

ticipants pressed the onscreen HUNT button and received feedback in terms of calories (1–1000).

Phase 2 consisted of 24 rounds of individual learning (simulating guided variation), in which participants could modify one or more characteristics of their points. During Phase 2 there was no opportunity to view or copy the point of any other participant, either from the pretest or the immediate group. Participants could change the values for one, more than one, or none of the five point attributes and at any time press the onscreen VIEW button, whereupon an image of the point with the appropriate length, color, and so on was displayed on the screen. When participants were happy with their designs, they pressed the onscreen HUNT button, whereupon feedback was given in terms of calories, as in Phase 1. During each hunt, participants could see their score from the previous hunt, their rank within the group (for participants in groups of four or more), their cumulative score (the running total of all hunt scores up to that hunt), and the number of hunts left in the season.

The final five hunts comprised Phase 3, representing horizontal cultural transmission, in which participants could select one of the other members of their group and view their point characteristics. Next to each group member's name was his or her cumulative score, allowing the participant to select the most successful group member. If the participant chose to view another participant's values, then those values appeared in red next to the participant's own values, which did not automatically change. At all times participants could modify their own values, as in Phase 2, whether they viewed another participant's values or not. If they did, participants were free to change one or more of their own values to match the other participant's values or ignore the other participant's values. Feedback was the same as for Phase 2.

Each participant played three seasons each comprising 30 hunts, with each season divided into the three phases of learning, except for the six pretest models (see above) and all participants in groups of size 1–3, who did not play Phase 3 as there were not enough other group members present. For these participants Phase 2 lasted from Hunt 2 to Hunt 30. For all participants, the cumulative score was reset to zero at the start of each season, and each season featured different optimal values, which constituted the selective environment. Optimal values

stayed constant within seasons. There were three different sets of optimal values (three different environments), to which participants were assigned at random. All participants in the same group experienced the same environment. During Season 3, participants additionally had to pay a cost to modify their point designs. During each hunt of Season 3 (except Hunt 1), participants could either keep the same values from the previous hunt at no cost or choose to modify their values, whereupon 250 calories were deducted from their cumulative score. Viewing another participant's point design in Phase 3 entailed no cost, but choosing to modify that point did entail the cost.

Procedure

Participants completed the experiment in groups of 1–6, depending on availability. Each participant sat at a separate computer terminal. Dividers ensured that participants could not see the screen of adjacent participants. On-screen instructions explained the experimental design and provided instructions. Participants were instructed not to communicate with each other except by means of the program, or to write anything down.

Fitness Calculations

Following each hunt, participants were given feedback in terms of the number of calories (1–1,000) that their current point design yielded. This feedback was calculated by comparing the participants' values of each of their point attributes with corresponding hidden optimal values. The closer a participant's values of each attribute were to the corresponding optimal value for that attribute, the higher the overall feedback score. Example optimal fitness functions are shown in Figure 1. Each environment (E0, E1, E2, and E3) had three different sets of optimal values, one for each season of hunting. These optimal values stayed constant during each season and changed at the beginning of each season. The overall feedback is denoted here by W ($0 < W \leq 1,000$) and was calculated from four of the five attributes, Length, Width, Thickness, and Shape (color did not affect point performance and therefore constitutes a neutral trait). The four functional attributes contributed independently to the overall fitness of a point design, W , differentially weighted according to equation (1):

$$W = 1000 (.275 W_\lambda + .25 W_w + .35 W_t + .125 W_s) \quad (1)$$

where subscripts λ , w , t , and s denote the attributes Length, Weight, Thickness, and Shape respectively (λ is used for Length instead of lower-case "l" to avoid confusion with numeral "1"). Hence, W_λ ($0 < W_\lambda < 1$), W_w ($0 < W_w < 1$), W_t ($0 < W_t < 1$), and W_s ($0 < W_s < 1$) are the separate fitness contributions of Length, Width, Thickness, and Shape, respectively. Note that participants received only the single overall number, W , and were not informed of the separate contributions to this number from each attribute. Equation 1 states that Thickness has the greatest effect on the overall feedback, followed by Weight, then Length, and finally Shape. We can therefore predict that the optimal strategy is to modify attributes in this order of priority.

The three continuous variables, Length, Width, and Thickness, each had bimodal fitness distributions. Fitness was calculated from two normally distributed functions for each attribute, W_1 and W_2 , where W_1 is centered around a global optimum value, O_1 ($10 < O_1 < 90$), and W_2 is centered around a local optimum value O_2 ($10 < O_2 < 90$). If the participant's attribute value is at the global optimum, O_1 , then the participant receives the maximum possible fitness for that attribute ($W_\lambda = 1$, $W_w = 1$, or $W_t = 1$). If the participant's value is at the local optimum, O_2 , then the participant receives two-thirds of the fitness of the global optimum ($W_\lambda = .66$, $W_w = .66$, or $W_t = .66$). Any deviation from either of these optima decreases the feedback from that attributes' fitness and hence the overall feedback, W , according to equation (1). Equations (2) and (3) give functions W_1 and W_2 for Length (hence $W_{\lambda 1}$ and $W_{\lambda 2}$), adapted from Boyd and Richerson (1985:100):

$$W_{\lambda 1}(X_\lambda, O_{\lambda 1}) = P_1 \exp [-(X_\lambda/100 - O_{\lambda 1}/100)^2 / (2s)] \quad (2)$$

$$W_{\lambda 2}(X_\lambda, O_{\lambda 2}) = P_2 \exp [-(X_\lambda/100 - O_{\lambda 2}/100)^2 / (2s)] \quad (3)$$

where X_λ is the participant's Length value; $O_{\lambda 1}$ and $O_{\lambda 2}$ are the specific values of the two optima in terms of the arbitrary Length units; P_1 and P_2 are the maximum fitnesses given by the global and local optima, respectively; and s is a measure of the intensity of selection (the shallowness of the curve in Figure 1). In this study P_1 , P_2 , and s were con-

stant at $P_1 = 1$, $P_2 = .66$, $s = .025$. The overall W_λ is then given by the larger of the two values, $W_{\lambda 1}$ and $W_{\lambda 2}$ (equation 4):

$$W_\lambda = \begin{cases} W_{\lambda 1} & \text{if } W_{\lambda 1} > W_{\lambda 2} \\ W_{\lambda 2} & \text{if } W_{\lambda 2} > W_{\lambda 1} \end{cases} \quad (4)$$

Equivalent variables for Weight and Thickness can be substituted into these equations to give similar fitness functions but at different optimal values, as shown in Figure 1. The discrete attribute, Shape, could take one of four values (S1, S2, S3 or S4). These shapes were assigned a rank order that determined their fitness payoff, W_s (Figure 1). The first shape gave the maximum $W_s = 1$, the second gave $W_s = .9$, the third $W_s = .66$, and the last $W_s = .33$. So, for example, if the rank order was {S3, S2, S4, S1}, then a choice of S3 gave $W_s = 1$; S2 gave $W_s = .9$; S3 gave $W_s = .66$; and S1 gave $W_s = .33$. Finally, the feedback shown to the participants was subject to random error (as in McElreath et al. 2005), intended to represent random factors such as weather fluctuation or prey availability that affect feedback independently of point design. The fitness shown to the participants, W' ($0 < W' \leq 1,000$), was randomly taken from a normal distribution, with mean W and standard deviation E . Parameter E therefore represents the magnitude of the error, here constant at $E = 5$.

Statistical Analyses

As a result of the bimodal fitness functions, many of the dependent variables exhibited a bimodal rather than a normal distribution, so nonparametric correlation statistics are used to compare the point attributes. Correlations between two quantitative variables (Length, Width, Thickness) are Spearman's r_s ; correlations between one quantitative (Length, Width, Thickness) and one nominal (Shape, Color) variable are *Eta*; correlations between two nominal variables (Shape and Color) are Cramer's V . Significance of *Eta* was derived using Kruskal-Wallis tests. Regression models were not fitted because of nonindependence of data; each participant's attribute or score value is dependent on both that participant's previous values and, during Phase 3, on the values of other participants in the group. Dunn and Clark's z -test was used to test the hypotheses that correlations between the

same attributes changed significantly between different experimental phases, as specified by Chen and Popovich (2002:24–25).

Results

The following subsections address, in turn, each main research question addressed in the experimental simulation. Each subsection begins with a one-line summary of our findings for non-statistically inclined readers, followed by details of the statistical analyses used to support that conclusion.

Did Participants Copy the Most Successful Individuals in Phase 1?

Summary: Yes, the majority of participants in Phase 1 chose to copy the single most-successful model, consistent with the operation of indirectly biased oblique cultural transmission.

Details: In Phase 1 (Hunt 1) participants were presented with a list of the six pretest models and the models' final scores, and they could select one of those six models to copy (simulating oblique cultural transmission). We predicted that participants would choose to copy the model with the highest score—they would exhibit indirect bias. In all three seasons the majority (but by no means all) of participants acted in accordance with this prediction: In Season 1, 66/75 (88.0 percent) of participants chose the most successful model to copy; in Season 2, 57/75 (76.0 percent); and in Season 3, 59/75 (78.7 percent). These choices significantly differed from random according to chi-square tests (Season 1: $\chi^2(5) = 275.96$, $p < .001$; Season 2: $\chi^2(5) = 193.08$, $p < .001$; Season 3: $\chi^2(5) = 212.12$, $p < .001$).

Were the Point Attributes More Highly Correlated Following Phase 1 (Indirect Bias) Than Following Phase 2 (Guided Variation)?

Summary: Yes, point attributes were more highly correlated following Phase 1 than following Phase 2, consistent with the operation of indirectly biased oblique cultural transmission, and re-creating the patterns documented by Bettinger and Eerkens (1999) for prehistoric Great Basin points.

Details: As argued by Bettinger and Eerkens (1999), indirectly biased cultural transmission should produce strong and significant correlations

between point attributes. Hence, we should see stronger correlations following Phase 1 (Hunt 1), having established above that the majority of participants were indeed copying the single-most successful model, than following Phase 2 (Hunt 25), after 24 hunts of individual trial and error learning in different environments and in a multimodal adaptive landscape. Table 3 shows the 10 correlations between the point attributes at these two hunts (the upper two correlations in each cell) separately for each season. At Hunt 1, for Season 1 all 10 correlations were highly significant at $p < .001$; for Season 2, eight of the 10 were significant at $p < .001$, one at $p < .01$, and one at $p < .05$; for Season 3 seven were significant at $p < .001$, one at $p < .01$, and the other two were not significant ($p > .05$). At Hunt 25, only two correlations were significant at $p < .05$ for Season 1, one was significant for Season 2, and five were significant for Season 3. However, given that we are performing multiple simultaneous statistical tests, we must adjust our significance level to reflect the increased probability of false positives. Adopting a Bonferroni-corrected significance level for 10 comparisons of $\alpha^* = \alpha / N = .05 / 10 = .005$ (where N = number of comparisons) means that none of the Hunt 25 correlations are significant, whereas the majority of the Hunt 1 correlations (25/30) remain significant.

We can also test whether the correlations compare in the directions predicted, that is, whether $\text{corr}_{H1} > \text{corr}_{H25}$. From Table 3, 10/10 comparisons from Season 1, 10/10 from Season 2, and 8/10 from Season 3 fulfill this relationship. As shown in Table 4, z-tests showed that every attribute comparison for every season was correlated significantly less strongly at Hunt 25 than at Hunt 1, as predicted (except for two anomalous comparisons in Season 3, Thickness x Shape and Thickness x Width, which were significant in the opposite direction).

Was the Phase 2 Guided-Variation Divergence a Result of the Different Environments or of the Multimodal Adaptive Landscape?

Summary: The multimodal adaptive landscape at least partly caused the divergence due to guided variation/individual learning during Phase 2.

Details: The divergence resulting from individual trial-and-error learning in Phase 2, which produced low and predominantly nonsignificant attribute correlations, can be attributed to (1) the

Table 3. Correlation Coefficients for Experimental Point Attributes, by Learning Phase (Social or Individual).

	Width			Thickness			Shape			Color		
	S1			S2			S3			S1		
	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
Height	Hunt 1 (SL)	.801***	-.916***	-.846***	.996***	.243*	.339**	1.000***	.680***	.952***	.828***	.964***
	Hunt 25 (IL)	.272*	-.179	-.331**	-.004	-.093	-.160	.248	.450**	.281	.187	.159
	Hunt 30 (SL)	-.432	.192	-.864***	-.280	-.397	-.502*	.730*	.717	.636	.664*	.582
Width	Hunt 1 (SL)				.752***	-.327**	-.210	.856***	.478***	.945***	.994***	.997***
	Hunt 25 (IL)				-.014	.072	.285*	.300	.170	.422**	.350*	.323
	Hunt 30 (SL)				-.515*	-.229	.366	.763*	.505	.345	.342	.598
Thickness								.895***	.809***	.009	.706***	.555***
								.162	.165	.381**	.190	.152
								.867**	.466	.766**	.489	.846
Shape												
Color												
Hunt 1 (SL)												
Hunt 25 (IL)												
Hunt 30 (SL)												

Note: Correlations are Spearman's r_s . S1 = Season 1; S2 = Season 2; S3 = Season 3. SL = Social Learning; IL = Individual Learning. Sample sizes: Hunts 1 and 25, N = 75; Hunt 30, N = 18.
* $p < .05$; ** $p < .01$; *** $p < .001$.

Table 4. Z Scores for the Comparison between Correlations at Hunt 1 and Hunt 25.

	Width			Thickness			Shape			Color		
	S1			S2			S3			S1		
	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
Height	5.00***	-.8.32***	-.5.32***	18.55***	2.11*	3.05**	28.88***	2.15*	9.45***	6.10***	3.40***	11.49***
				5.92***	-.2.50**	-.2.99***	5.90***	2.12*	8.52***	15.35***	2.85*	18.52***
Width												
Thickness												
Shape												
Color												
Hunt 1 (SL)												
Hunt 25 (IL)												
Hunt 30 (SL)												

Note: All correlations are larger at Hunt 1 than at Hunt 25, except those marked ^a (correlation larger at Hunt 25).
* $p < .05$; ** $p < .01$; *** $p < .001$.

fact that participants/groups were split into three different environments (sets of optimal values), and/or (2) the fact that those environments constituted adaptive landscapes with multiple locally optimal fitness peaks. Although the study was not specifically designed to address this, we can get some idea by comparing the correlations shown in Table 3 with their equivalent partial correlation coefficients controlling for environment. If the low correlations at Hunt 25 are the result of the different environments, then removing the effect of environment should increase the correlations. If the multimodal adaptive landscape is responsible, then the correlations should remain low. Partial correlations support the latter, with the majority of the correlations showing little change after removing the effects of environment. Only two of the 30 correlations changed in significance, with length \times width for Season 2 ($r_p = -.233, p < .05$) and width \times color for Season 3 ($r_p = -.265, p < .05$) becoming significant. Overall, however, the similarly low and predominantly non-significant partial correlations suggest that the multimodal adaptive landscape at least partly caused the divergence.

Did Participants View the Most Successful Individuals' Designs in Phase 3 (Horizontal Cultural Transmission)?

Summary: Yes, participants during Phase 3 preferentially accessed the point design of the most successful group member, consistent with the operation of indirectly biased horizontal cultural transmission.

Details: During Phase 3 (Hunts 26–30), participants in groups of 4–6 could choose to view and copy the point design of another group member, i.e., engage in horizontal cultural transmission. Participants could see the other players' cumulative scores, and we again predicted that participants would be more likely to view the point attributes of the most successful group member. We excluded from these analyses the most successful player in each group in each season, as the most successful player cannot copy him- or herself (the top player often changed between seasons, so typically a different player is excluded in each season). Of the remaining 29 participants, the majority viewed at least one other player's point values over the three seasons: 27/29 (93.1 percent) during Season 1, 26/29 (89.7 percent) during Season 2, and 25/29

(86.2 percent) during Season 3.

Despite being allowed to view another player for free on each of the last five hunts of each season, the mean number of times that another point was accessed per participant was much lower than the potential maximum of five (1.72 for Season 1, 1.59 for Season 2, and 1.21 for Season 3). The mean number of different players accessed was also low, with a mean of around one (1.17 for Season 1, 1.14 for Season 2, and .96 for Season 3). This single model was typically the most successful: of those participants who accessed at least one other player's point design, 25/27 (92.6 percent) viewed the most successful player (the player with the highest cumulative score at Hunt 25) at least once during Season 1, 21/26 (80.8 percent) during Season 2, and 22/25 (88.0 percent) during Season 3. These copying choices are significantly different from those expected if participants were viewing players at random (Season 1: $\chi^2(4) = 63.93, p < .001$; Season 2: $\chi^2(4) = 60.92, p < .001$; Season 3: $\chi^2(4) = 69.75, p < .001$).

Were the Point Attributes More Highly Correlated Following Phase 3 (Horizontal Cultural Transmission) Than Following Phase 2 (Guided Variation)?

Summary: A largely nonsignificant trend suggests that point attributes were more highly correlated following Phase 3 than following Phase 2, consistent with the operation of indirectly biased horizontal cultural transmission.

Details: We again predicted stronger correlations following the horizontal cultural transmission of Phase 3 (at Hunt 30) than following Phase 2 (at Hunt 25), after 24 hunts of individual trial-and-error learning. The most appropriate data here were obtained from the three six-person groups, each of which was assigned to a different one of the three environments, giving equal sample sizes for each environment and, as these groups were largest, the best chance to detect a significant effect of horizontal cultural transmission. Table 3 shows the 10 correlations between the point dimensions for these two hunts (the lower two correlations in each cell). For Season 1, 9/10 of the correlations showed an increase from Hunt 25 to Hunt 30 (i.e., $\text{corr}_{H30} > \text{corr}_{H25}$), as predicted. Six of these were significant at $p < .05$ for Hunt 30, compared to two for Hunt 25. Nine of the Season 2 correlations

increased from Hunt 25 to Hunt 30, although only one Hunt 30 correlation was significant. For Season 3, 9/10 correlations again showed an increase, although the number of significant correlations for Hunt 30—three—actually fell from the five significant for Hunt 25. However, comparisons here may be less valid as a result of the unequal sample sizes for Hunt 30 ($n = 18$) and Hunt 25 ($n = 75$), which also prevents the use of z -tests, whereas adopting a Bonferroni-corrected α means that only one of the Hunt 30 correlations is significant. Nevertheless, we note the general pattern whereby almost all attribute comparisons in all seasons showed an increased correlation coefficient at Hunt 30 following indirectly biased horizontal cultural transmission compared with Hunt 25 following individual trial-and-error learning.

Was There Less Within-Group Variation as a Result of Cultural Transmission?

Summary: Yes, there was significantly less within-group variation during periods of cultural transmission compared with periods of individual learning, consistent with the homogenizing effect of indirectly biased within-group horizontal cultural transmission.

Details: Figures 2a, 2b, and 2c show for Seasons 1, 2, and 3, respectively, the within-group coefficient of variation (WGCV) and between-group coefficient of variation (BGCV) for length, width, and thickness combined (the mean CV for all three measures), for the three six-person groups. A low WGCV indicates homogeneity within groups. This could be because of either (1) individual learning—different groups are learning in different environments, or sets of optimal values, so individual learning would be expected to cause each group to converge on their different optima, or (2) horizontal cultural transmission—following Hunt 25, group members are copying the single most successful player in their group, so groups become homogenized. Figures 2a and 2b suggest that (2) is operating mainly during Seasons 1 and 2, given that marked differences between WGCV and BGCV occur only (and abruptly) following Hunt 25, when cultural transmission within the groups is permitted. These differences in variance were tested using Kruskal-Wallis one-way ANOVAs. At Hunt 30, following cultural transmission, there was a significant effect of group for

length, width, and thickness for Season 1 (length: $\chi^2(2) = 15.20$, $p < .001$; width: $\chi^2(2) = 15.46$, $p < .001$; thickness: $\chi^2(2) = 12.42$, $p < .002$), Season 2 (length: $\chi^2(2) = 12.00$, $p < .002$; width: $\chi^2(2) = 11.83$, $p < .002$; thickness: $\chi^2(2) = 6.79$, $p < .05$), and Season 3 (length: $\chi^2(2) = 15.63$, $p < .001$; width: $\chi^2(2) = 13.64$, $p < .001$; thickness: $\chi^2(2) = 11.00$, $p < .004$), indicating the homogenizing effect of within-group horizontal cultural transmission. At Hunt 25, the last individual-learning trial, Season 1 showed no effect of group for any of the three attributes (length: $\chi^2(2) = 5.56$, ns; width: $\chi^2(2) = 5.96$, ns; thickness: $\chi^2(2) = .90$, ns); Season 2 showed an effect of group for one of the attributes (length: $\chi^2(2) = 12.74$, $p < .002$; width: $\chi^2(2) = 2.11$, ns; thickness: $\chi^2(2) = 1.17$, ns); and Season 3 showed an effect of group for two of the attributes (length: $\chi^2(2) = 12.89$, $p < .002$; width: $\chi^2(2) = 11.83$, $p < .003$; thickness: $\chi^2(2) = 4.86$, ns). Figure 2 also shows that there was less variation both between and within groups during Season 3 than during Seasons 1 and 2, a result of the cost of modification. This cost prevented participants from experimenting with their point designs, reducing the divergence and also attenuating the homogenizing effect of horizontal transmission in Phase 3, given that the groups were already homogenized, as suggested by the Season 3 ANOVAs.

Did the Groups Outperform the Individual Controls When They Had Access to Social Information?

Summary: Yes, groups significantly outperformed individual controls when the groups were permitted to engage in cultural transmission, indicating that indirectly biased cultural transmission was relatively more adaptive than individual learning.

Details: Figures 3a, 3b, and 3c show the mean scores during Seasons 1, 2, and 3, respectively, of the 27 individual controls and the 18 participants in groups of six, the latter of which could engage in horizontal cultural transmission during the final five hunts. In all three seasons the groups have higher mean scores than the individuals during those final five hunts, as predicted. Mixed ANOVAs with hunt as a within-group factor and group size ($n = 1$ vs. $n = 6$) as a between-groups factor showed, for the last five hunts (Hunts 26–30), a significant effect of group size for all three seasons and a larger effect size for Season 3, consistent with the

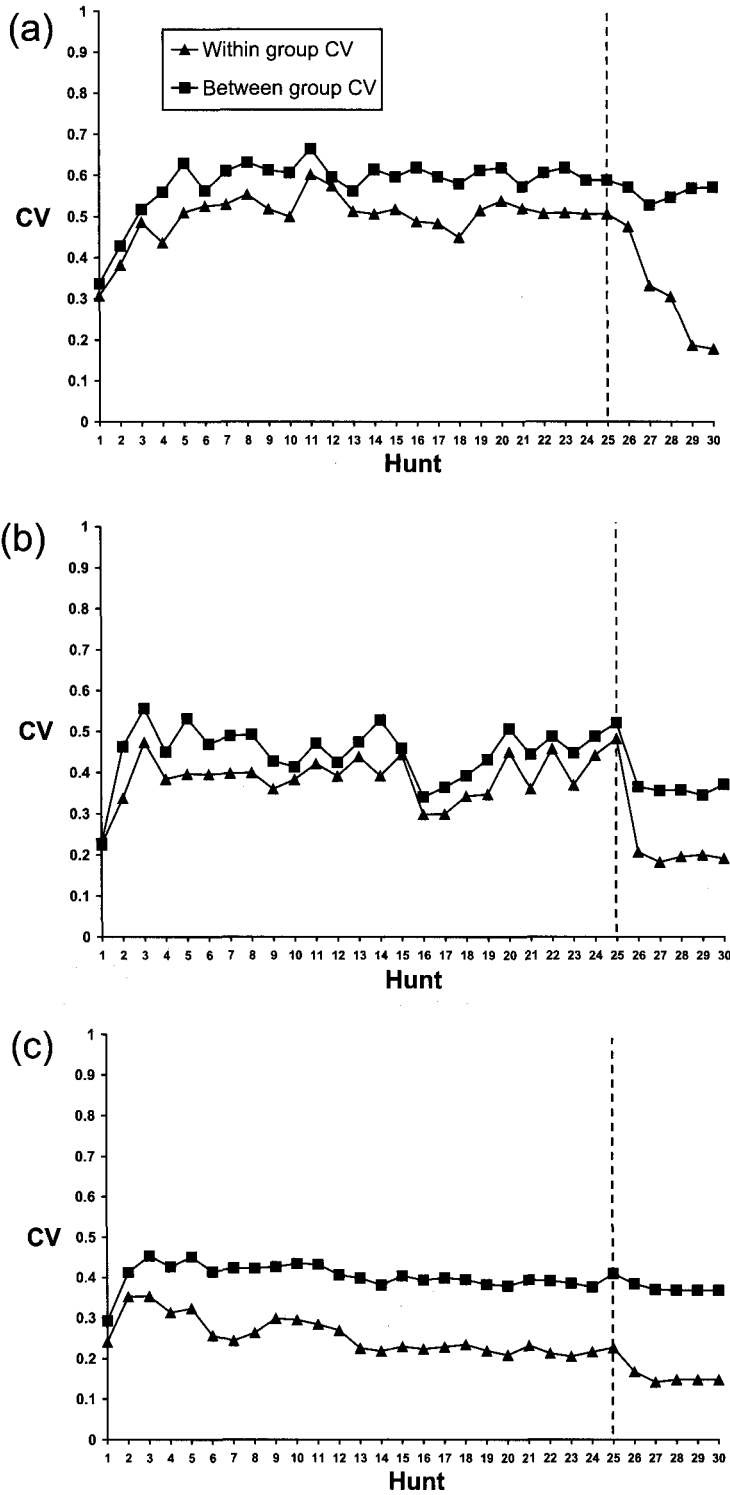


Figure 2. Within-group and between-group coefficient of variation (CV) for the six-person groups (N = 18) for (a) Season 1, (b) Season 2 and (c) Season 3. Values plotted are the mean CV for Height, Width and Thickness. The dashed vertical line marks the final individual learning hunt before social learning was permitted.

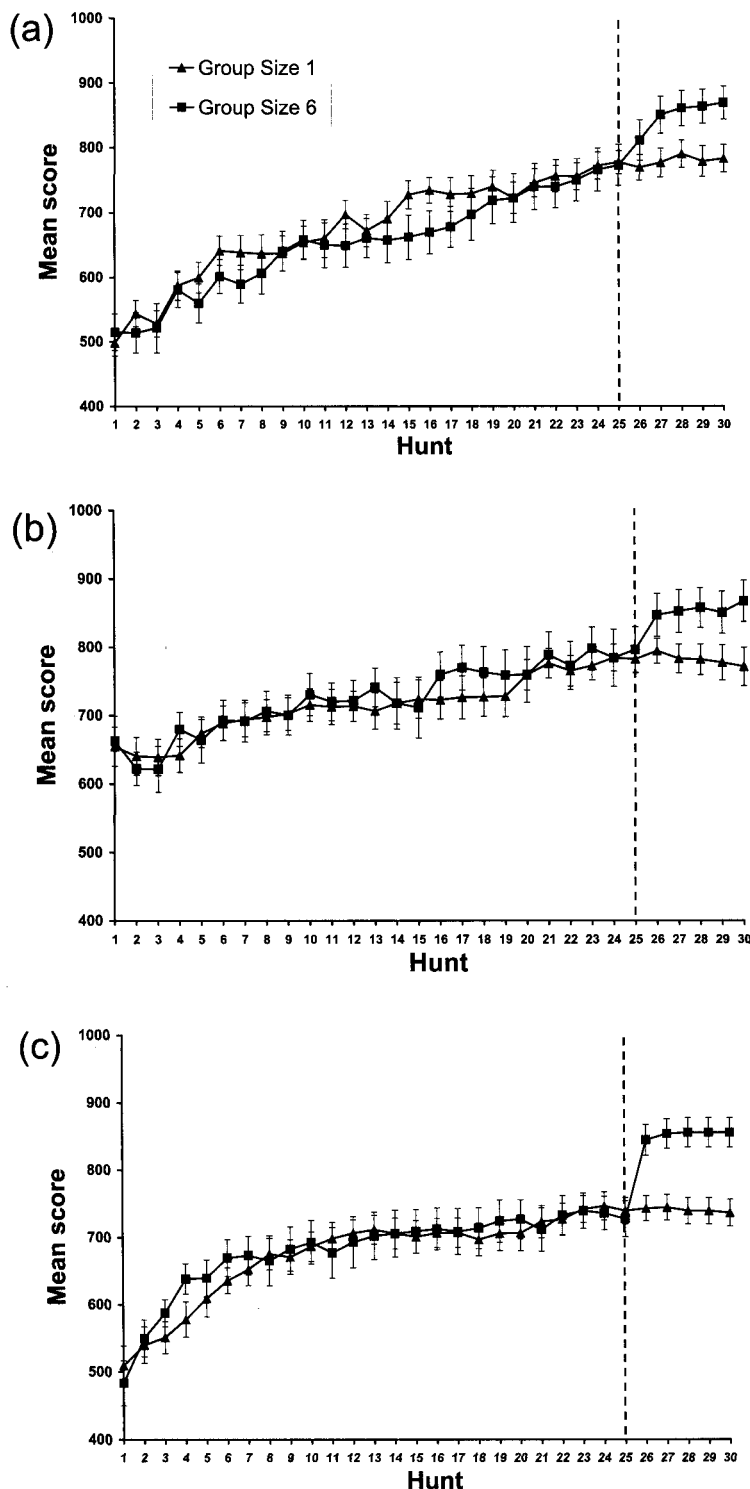


Figure 3. Mean participant score during (a) Season 1, (b) Season 2 and (c) Season 3 of individual controls ($N = 27$) and participants in groups of six ($N = 18$). The dashed vertical line marks the final individual learning hunt before social learning is permitted in the groups. Error bars show standard errors.

increased cost of individual learning (Season 1: $F(1,43) = 4.76, p < .05, \eta^2 = .10$; Season 2: $F(1,43) = 4.23, p < .05, \eta^2 = .09$; Season 3: $F(1,43) = 14.85, p < .001, \eta^2 = .27$). ANOVAs on the previous five hunts (Hunts 21–25), during which the groups could not access social information, showed no effect of group size—Season 1: $F(1,43) = .05, ns$; Season 2: $F(1,43) = .10, ns$; Season 3: $F(1,43) = .03, ns$.

Did Group Participants Benefit from Social Information by Jumping to Globally Optimal Peaks?

Summary: Yes, participants in groups, but not individual controls, showed a significant tendency to jump from locally optimal peaks to globally optimal peaks in the fitness landscape, as a result of indirectly biased cultural transmission.

Details: Here we calculated, for each participant and for each season, the number of continuous attributes (length, width, and thickness) that were within five units either side of that attributes' global optimum ($O_1 \pm 5$), giving a number between zero and three. We predicted that this number would be greater at Hunt 30 (after cultural transmission) than at Hunt 25 (before cultural transmission) for the participants in groups of six. Wilcoxon's signed-ranks tests showed this difference to be significant for all three seasons—Season 1: $z = 2.71, p < .01$; Season 2: $z = 2.27, p < .05$; Season 3: $z = 2.57, p < .01$. For the individual control participants, this comparison was not significant for Season 2 ($z = 1.41, ns$) or Season 3 ($z = 1.00, ns$), whereas Season 1 showed a significant decrease ($z = -2.45, p < .05$). It is unclear why individual-learning controls showed this decrease, but the hypothesis that indirectly biased cultural transmission increases the number of participants at global optima is supported and cannot be attributed simply to individual improvement over time. We also tested the specific hypothesis that access to social information allows participants to jump from locally to globally optimal fitness peaks. Hence we calculated the number of continuous attributes around the local optimum ($O_2 \pm 5$), and compared this value at Hunt 25 with the number of attributes within five units of the globally optimal peak at Hunt 30, predicting an increase for the group participants and no difference for the individual participants. As predicted, Wilcoxon's signed-ranks test showed

significant increases for the group participants for all three seasons—Season 1: $z = 3.40, p < .001$; Season 2: $z = 2.18, p < .05$; Season 3: $z = 2.85, p < .01$. For the individual participants, there was no significant increase during any season—Season 1: $z = 1.80, ns$; Season 2: $z = .21, ns$; Season 3: $z = 1.52, ns$.

How Did Individuals Learn on Their Own?

Summary: Most individuals changed one point attribute per hunt by a magnitude of five units, except when their score was low (when they changed more attributes by greater magnitudes) and when individual learning was costly (when they changed attributes less often by greater magnitudes). Continuous attributes (e.g., length) were modified more frequently than discrete attributes (e.g., shape).

Details: As well as exploring the effects of cultural transmission, we can also look at how individuals engaged in the task on their own (both the individual participants and the group participants during Phase 2). A player's strategy can be defined usefully by two parameters. During each hunt, a participant can choose to modify one, more than one, or none of the five attributes. The parameter, d ($0 \leq d \leq 5$), is defined as the number of attributes changed during one hunt. The second parameter, c ($0 \leq c \leq 99$), is defined as the mean amount by which the participant changed the continuous attributes (Length, Width, and Thickness), ranging from 0 (no change) to 99 (largest possible change). Both d and c are integers and take no value for Hunt 1 when there was no modification. Figure 4a shows that d rapidly fell during the first seven hunts of each season before leveling off, apart from an increase during Hunt 26, when some of the participants copied the points of other players (especially in Season 3). Excluding Hunts 1–6 and Hunt 26, the majority of participants on the majority of hunts chose to change one dimension ($d = 1$) during Season 1 (mean = 1.43, median = 1, mode = 1) and Season 2 (mean = 1.22, median = 1, mode = 1), and most chose not to change any dimension ($d = 0$) during Season 3 (mean = .45, median = 0, mode = 0). There was also a significant negative correlation between Score (number of calories received during a single hunt) and d (Season 1: $r_s = -.368, p < .01$; Season 2: $r_s = -.369, p < .01$; Season 3: $r_s = -.243, p < .01$), suggesting that players respond to

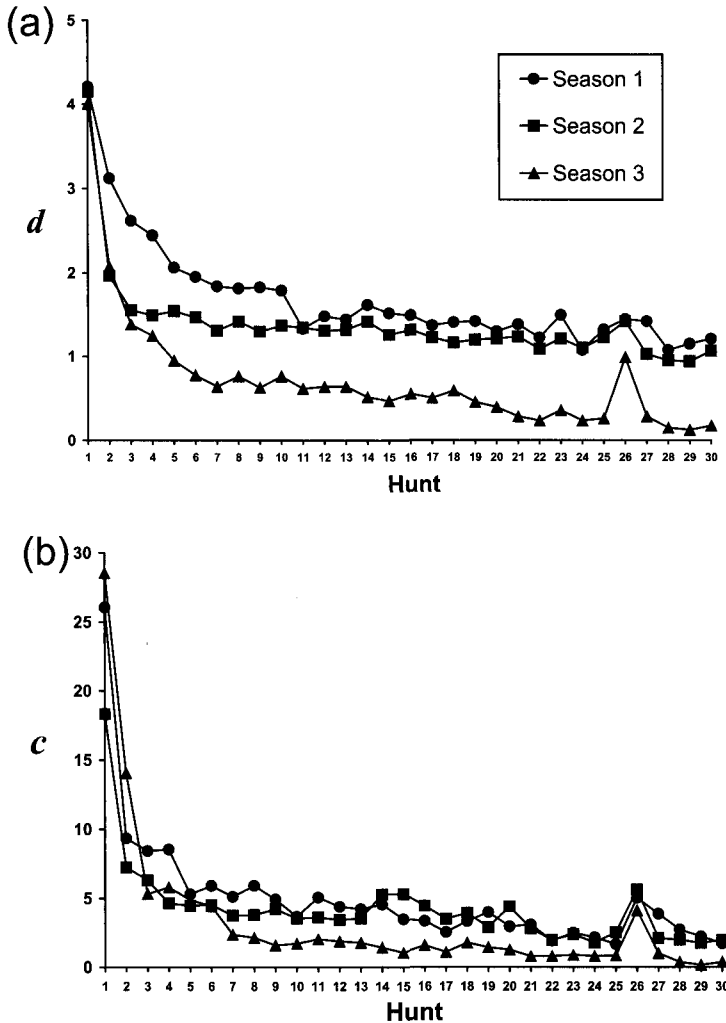


Figure 4. (a) Mean number of dimensions changed per hunt (parameter d), for each season. Data from all 81 participants. (b) Mean unit change in the continuous dimensions (Height, Width and Thickness combined) per hunt (parameter c), for each season. Data from all 81 participants, and excluding cases where the dimension was not changed.

a low score by changing more dimensions at once. Figure 4b shows a similar pattern of change for c . Excluding Hunts 1–6 and Hunt 26, most participants during the first two seasons made changes of five units (Season 1: mean = 9.17, median = 5, mode = 5; Season 2: mean = 9.50, median = 5, mode = 5). In Season 3 this increased to 10 units (Season 3: mean = 10.59, median = 10, mode = 10) as a result of the increased cost of individual learning. There was a significant negative correlation between score and c (Season 1: $r_s = -.323$, $p < .01$; Season 2: $r_s = -.423$, $p < .01$; Season 3: $r_s = -.304$, $p < .01$), suggesting that players respond to a low

score by increasing the amount by which they change the dimensions.

Figures 5a, 5b, and 5c show for each hunt the proportion of the total 81 participants who changed each dimension, for Seasons 1, 2, and 3 respectively. for length, width, and thickness, the frequency of modification started around .6 to .85 for Season 1 before leveling off at around .35 (mean for Hunts 7–30: length = .34, width = .34, thickness = .35). Season 2 continued this trend (mean for Hunts 7–30: length = .33, width = .33, thickness = .33), whereas Season 3 featured lower frequencies as a result of the cost of modification

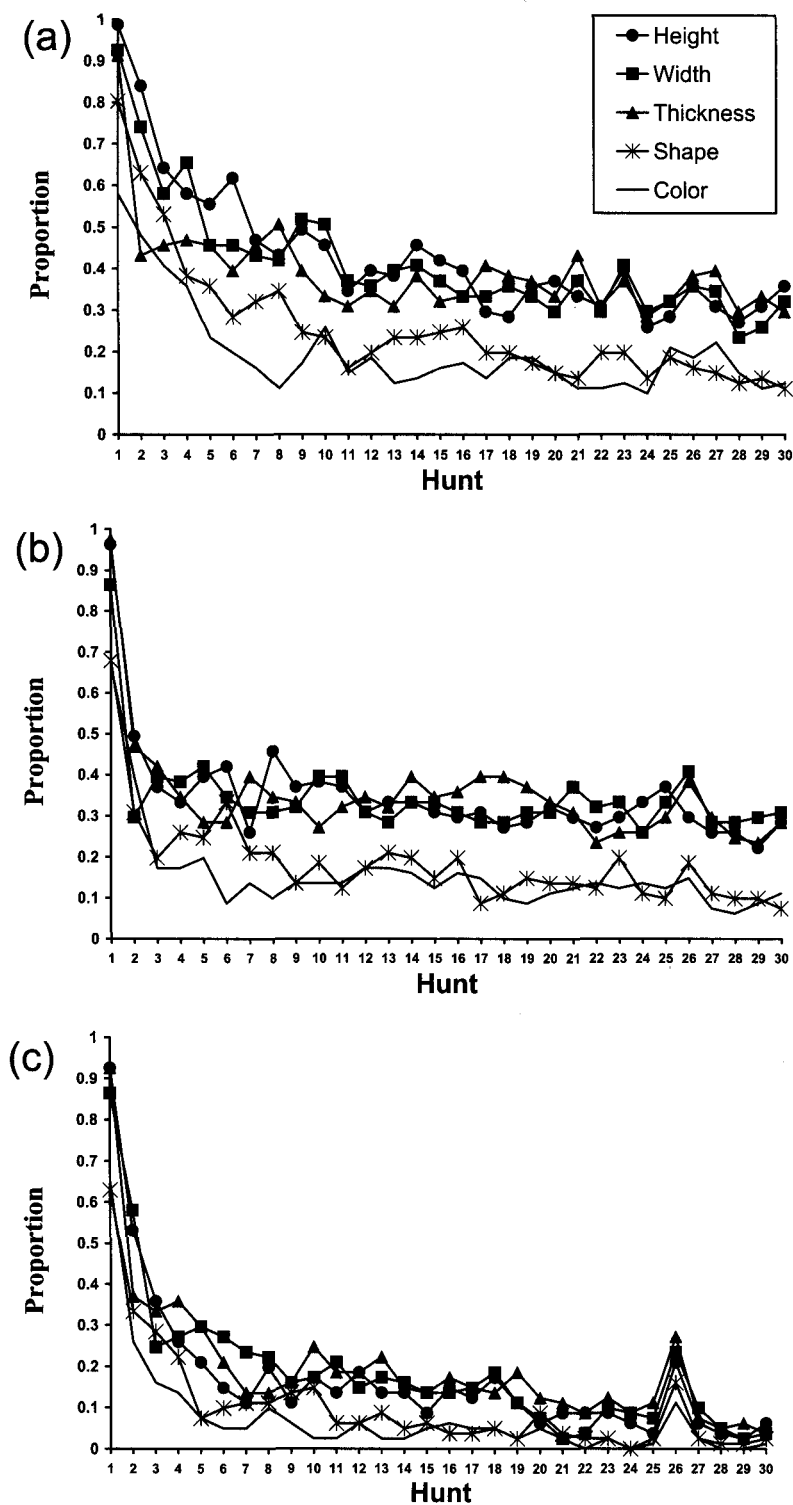


Figure 5. Proportion of participants during (a) Season 1, (b) Season 2 and (c) Season 3 who, on each hunt, changed each dimension.

Table 5. Paired T-Tests Comparing the Frequency with Which Each Attribute Was Modified.

Comparison	Season 1	Season 2	Season 3
Height x Width	.49	-.70	-2.06
Height x Thickness	.41	-.70	-3.85**
Width x Thickness	.12	-.16	-1.29
Height x Shape	16.63***	14.62***	6.08***
Height x Color	14.27***	17.92***	9.92***
Width x Shape	14.70***	17.16***	7.18***
Width x Color	15.21***	22.75***	8.80***
Thickness x Shape	14.59***	15.57***	9.51***
Thickness x Color	14.57***	18.82***	10.36***
Shape x Color	2.96**	2.63**	3.41**

Note: Values shown are *t* values with *df* = 23.

p* < .05; *p* < .01; ****p* < .001.

(mean for Hunts 7–30: length = .14, width = .16, thickness = .17). Shape and color were changed less frequently than length, width, and thickness in Season 1 (mean for Hunts 7–30: shape = .18, color = .15), Season 2 (mean for Hunts 7–30: shape = .17, color = .14), and Season 3 (mean for Hunts 7–30: shape = .08, color = .05). Paired *t*-tests on these frequencies (Table 5) from Hunts 7–30 confirmed that participants showed a preference for changing the three continuous dimensions over the two discrete dimensions. No differences were found for Seasons 1 and 2 between the continuous attributes (length vs. width vs. thickness), although thickness was modified more than length during Season 3. The greater fitness contribution of Thickness (equation 1) may therefore have had an effect when participants became more experienced with the environment, and when modification was costly and there was a greater incentive to improve point fitness. Shape also showed significantly more modification than color, perhaps reflecting the fact that color is neutral.⁶

Discussion

The primary aim of this study was to experimentally simulate the two different transmission processes—guided variation and indirect bias—that Bettinger and Eerkens (1999) hypothesized to have generated differences in the degree to which arrow projectile-point attributes are correlated with each other in two regions of the Great Basin. They argued that points from eastern California, for which attributes such as length or width were poorly correlated, originally spread through a process of guided variation, where individuals first adopted

point technology from a model, then experimented with that design using individual trial and error. Hence, point designs diverged and correlations between point attributes fell. Points from central Nevada, meanwhile, feature highly correlated attributes, which Bettinger and Eerkens (1999) argued was because these points originally spread through a process of indirect bias, where individuals adopt point designs wholesale from a single successful model and do not modify or experiment with that design. Hence, every individual acquires and keeps the same point design, and the attributes become highly correlated.

Here, we simulated these two processes experimentally by (1) having participants copy the design of a model after being given information regarding that model's prior success (permitting indirect bias) and then (2) allowing the participants to experiment with their point designs in novel selective environments (permitting guided variation). The results matched the patterns of attribute correlation found by Bettinger and Eerkens (1999), with the former points exhibiting highly correlated attributes and the latter points exhibiting less correlated attributes. Hence, we can be more confident in Bettinger and Eerkens's (1999) hypothesis that these different archaeological patterns are the result of differences in how projectile-point technology in the two regions was originally transmitted.

To some extent these findings are unsurprising; during Phase 1 participants had no choice but to copy, and were only given information about success, so the obvious and straightforward choice was to copy the most successful model. We set the experiment up in this way primarily because previous experiments have found that participants tend

not to engage in social learning, even when it is optimal for them to do so (e.g., McElreath et al. 2005), and we wanted ensure that we had data that had been generated by social learning. (Note, however, that a potential alternative strategy for our participants during Phase 1 was to disregard information concerning relative success and choose a model at random; indeed, not every participant chose the most successful model during Phase 1.) In Phase 3, therefore, we provided a more stringent test of our prediction. During Phase 3 participants could copy the point design of another member of their immediate group, given information about the cumulative success of those group members. Participants were not forced to copy, and could choose to continue engaging purely in individual learning. Yet a large majority of participants chose to copy the point design of the most successful participant in their group, i.e., exhibit indirectly biased horizontal cultural transmission, making the correlations between the different point attributes higher in Phase 3 (indirectly biased horizontal cultural transmission) than in Phase 2 (guided variation). In other words, our participants were quite willing to discard the point design that they had spent several trials modifying and improving, and replace it with a better point design from the most successful member of their group. The results of Phase 3, then, further increase our confidence in Bettinger and Eerkens's hypothesis.

During Phase 3 we also found that participants in groups, who could engage in cultural transmission, had significantly higher scores than individual controls, who could not engage in cultural transmission, indicating that indirectly biased cultural transmission was, for this task, relatively more adaptive than individual learning. This relative advantage increased during Season 3 when we introduced a cost to modification, consistent with previous theoretical analyses that show that social learning is relatively more adaptive when individual learning is costly (Boyd and Richerson 1989, 1995). This latter finding suggests a possible explanation for the original differences observed in the Great Basin: Perhaps the prehistoric Nevadan environment was harsher in some respect, imposing a greater cost on experimentation and necessitating a greater reliance on indirectly biased cultural transmission, hence reducing attribute correlations relative to California. This hypothesis might be tested

with further archaeological study.⁷

A secondary goal of the study was to simulate cultural transmission in a multimodal adaptive landscape. Whereas previous experiments (e.g., Kameda and Nakanishi 2002, 2003; McElreath et al. 2005) have used relatively simple learning tasks requiring the participant to select one of two options (e.g., crops or rabbit locations), here we used a more complex learning task involving multiple continuous and discrete, functional, and neutral attributes. Some of these attributes had bimodal fitness functions, thus giving multiple locally optimal designs of varying fitness, i.e., a multimodal adaptive landscape. The results of our experiment suggested that this multimodal adaptive landscape was instrumental in generating and maintaining diversity in the virtual-point designs during the period of individual learning (Phase 2), because different individuals converged by chance on different locally optimal peaks. We also found that this multimodal adaptive landscape allowed participants who engaged in indirectly biased horizontal cultural transmission, where individuals copy the design of the most successful person in their environment, to escape from these local optima and jump to the globally optimal peak (or at least the highest peak found by people in that group). That is, the multimodal adaptive landscape was instrumental in making indirect bias relatively more successful than individual learning, and potentially more successful than other social learning strategies, such as conformity.⁸

How realistic is this assumption of a multimodal adaptive landscape? Boyd and Richerson (1992) have argued that multimodal adaptive landscapes are likely to be common in cultural evolution and may significantly affect the historical trajectories of artifact lineages, just as population-genetic models suggest that multimodal adaptive landscapes have been important in biological evolution by guiding historical trajectories of biological lineages (Arnold et al. 2001; Lande 1986; Simpson 1944). Indeed, many problems and tasks faced by modern and prehistoric people would likely have had more than one solution, some better than others, but all better than nothing, and solutions are likely to represent compromises between multiple functions and requirements. With respect to projectile points, Cheshier and Kelly (2006) recently summarized experimental evidence for tradeoffs in

point designs among such factors as accuracy, range, killing power, and durability, stating, for example, that "thin, narrow points have greater penetrating power, but wide, thick points create a larger wound that bleeds more easily" (Cheshier and Kelly 2006:353). Such functional tradeoffs would potentially produce multiple locally optimal point designs, with, for example, one optimal design maximizing penetrating power and another maximizing bleeding.⁹ At a more general level, the bow-and-arrow and the spear-and-spearthrower represent two alternative solutions to the same problem (firing projectiles), with the bow and arrow apparently more effective, given that it replaced the spear thrower in most regions (Knecht 1997). This change is likely to have been the result of horizontal cultural transmission, with a jump from a lower peak (spear thrower) to a higher peak (bow and arrow) in the available design space of projectile technology. And adaptive landscapes are not restricted to material culture; Bettinger and Baumhoff (1982) argued that different Great Basin subsistence strategies represent different locally adaptive peaks in a multimodal adaptive landscape. The point we wish to emphasize is that it is only by formally testing assumptions regarding selective environments with experimental simulations, as was done here, as well as with mathematical models and computer simulations, that we can make quantitative predictions that can then be tested with archaeological data.

Several other findings are worth highlighting. First, we found that color, a neutral trait that had no effect on fitness/score, showed a virtually identical pattern of correlations as the other functional attributes, indicating that the participants readily copied color along with the others. This is an example of a neutral trait "piggybacking" on functional traits as a result of indirectly biased cultural transmission, as anticipated by Boyd and Richerson (1985). It would be instructive to see to what extent maladaptive traits (that impose a fitness cost) can similarly piggyback on functional traits. Second, we found that the continuous functional attributes (height, width and thickness) were modified more frequently than the discrete functional attribute (shape). This was probably because it was much easier to determine the optimal shape (try all four shapes and see which gave a higher fitness, requiring exactly four trials) than to determine the opti-

mal values of the continuous attributes (which, because they varied from 1–100, would likely require many more than four trials to find). This might potentially explain apparent cases of "punctuated equilibria" in the archaeological record—periods of gradual change constitute modification of continuous attributes such as length, and abrupt change constitutes a change in a discrete attribute, for example to a new shape or material (we thank J. Eerkens for pointing this out). Third, our finding that increasing the cost of modification increases the reliance on social learning suggests that complex technologies, which are costly to invent, learn, and maintain, should be more dependent on social learning than simpler technologies (see also Henrich 2004). Similarly, technologies that directly affect survival, such as those related to hunting, should be more dependent on social learning than technologies that bear less directly on survival, such as decorative artifacts, given that the cost of experimentation in the former is higher, where failure can mean starvation. Finally, although conformist transmission is commonly seen as generating the conditions for cultural group selection, i.e., low within-group variation and high between-group variation (e.g., Henrich and Boyd 1998), our results suggest that indirectly biased cultural transmission can potentially generate the same conditions. Future studies might simulate cultural group selection directly by allowing multiple groups to compete with one another, as in studies by Insko et al. (1980, 1983) and recent studies in experimental economics (e.g., Gurerk et al. 2006), while systematically controlling the movement of people and/or transmission of knowledge across groups. Although this might be technically unfeasible in the psychology lab, the Internet provides an ideal "virtual laboratory" for such studies. Indeed, thousands of people are already buying, selling, and using projectile technology, such as arrowheads, in *Massively Multiplayer Online Role-Playing Games*, potentially offering a mine of untapped data regarding the innovation, trade, and cultural transmission of artifacts and behavior.

There are several other ways in which this study could be extended. Cheshier and Kelly (2006) demonstrated experimentally that the attributes of actual projectile points are functionally linked, with durability determined by the thickness:length ratio, unlike in our experiment where the point attributes

independently contributed to overall fitness. It would be instructive to see how participants would deal with this more complex selective environment and whether the effects of cultural transmission remain the same. This is one of many parameters and assumptions that might be modified in further studies. E , the magnitude of the random error in the feedback received by participants, and s , the coefficient of selection in equations (2) and (3), can both be modified to change the difficulty of individual learning. Increasing the difficulty of individual learning is predicted to increase the frequency of social learning (Boyd and Richerson 1995), in the same way that increasing the cost of individual learning does, which during Season 3 here was found to increase the adaptive benefit of cultural transmission relative to individual learning. The ratio $P_1 : P_2$ can also be manipulated to change the relative height of the two fitness peaks in Figure 1 and hence the shape of the adaptive landscape.

As noted above, the cultural transmission employed by our participants was to a large degree constrained and forced by us, the experimenters. We divided individual and social learning into separate phases, forced participants to copy during Phase 1, and to some extent guided their decision as to whom to copy by only presenting information regarding relative success. Undoubtedly, reality is and was much more complex. Learning is not divided into separate social and nonsocial phases; rather, people can choose when to engage in individual learning and when to employ different social learning strategies (or employ both simultaneously), and they can choose a model based on multiple sources of information. There are also many factors that we omitted from our simulations yet would likely have affected projectile-point evolution, such as the process of manufacturing a point, the availability of raw materials, reworking, discard patterns, and the sociocultural significance of points. A simulation that is as complex and messy as reality, however, is more or less useless. Our intention with this initial experimental simulation was to ensure as much as possible that the participants exhibited the different transmission biases that our hypotheses required. Future studies might lift such constraints in a systematic manner, for example, by presenting participants with more information regarding potential models, such as aspects of appearance or personality, or arbitrary

social markers, to determine whether relative success is still used to select models. Allowing participants to flexibly engage in social and individual learning throughout the experiment might generate information on producer-scrouter dynamics similar to those observed by Kameda and Nakanishi (2002). Our participants also had no choice whether to allow another participant to copy their point designs, which is probably unrealistic, especially in the competitive task that was simulated here. Future studies might allow successful models to set a price for access to their point design, enabling a prestige-based "information market" (Henrich and Gil White 2001).

Some might argue that experiments with modern-day Western undergraduates can tell us little about the behavior of prehistoric hunter-gatherers, given the extensive differences in demographics, lifestyles, and environments of these two groups. While this is to some extent a valid criticism, it is somewhat unavoidable—barring the invention of a time machine, we cannot go back and directly study prehistoric patterns of cultural transmission. The problem might be partially reduced by using participants who are more similar to prehistoric hunter-gatherers than Western undergraduates, such as contemporary hunter-gatherers, along the lines of recent cross-cultural experiments regarding altruism (Henrich et al. 2004), although even contemporary hunter-gatherers likely differ in many ways from prehistoric Great Basin hunters-gatherers. Another difference is participant motivation. Whereas prehistoric hunters were hunting for their survival, our participants were taking part for \$8, raising the question of whether our participants were sufficiently motivated. Future studies might increase motivation by matching payment to performance in the experiment, as is commonly done in economics experiments (although, as shown by Camerer and Hogarth [1999], performance-related payment does not necessarily increase motivation or affect participants' performance), and/or increase the absolute value of payment. More generally, however, we stress that what experiments lack in external validity they make up for in *internal validity*—the ability to manipulate variables, control for confounding variables, assign participants to control and experimental groups, accurately record complete datasets, etc. The key point is that

different methodologies complement one another: experimental methods, which have low external validity and high internal validity, can be used *alongside* other methods, such as archaeological methods, which have high external validity and low internal validity, to provide a more complete and rigorous understanding of the past than any one of these methods alone (see Mesoudi 2007, 2008). Indeed, we hope that the present study has shown that experiments can be useful tools for simulating the population-level effects of different cultural transmission processes and for exploring the relative efficacy of those processes in a way that is often impossible with archaeological data alone. The findings from experiments can inform computer simulations and mathematical models, which can in turn verify and extend the experimental findings. The findings of both experiments and models then generate additional hypotheses that can be tested with further archaeological investigation, thus leading to a fuller understanding of past cultural change than any one method or approach can yield alone. Although reality is and was undoubtedly much more complex and messy than is assumed in our experiment, this kind of experimental simulation can be valuable as a means of testing specific archaeological hypotheses regarding cultural-transmission biases. We reiterate that the use of experimental simulations naturally follows from an evolutionary approach to cultural change, which emphasizes how individual-level differences in transmission can have major effects at the population level. Simple, highly controlled experimental simulations of biological inheritance and selection have been enormously useful in explicating the complexities of biological evolution, and experimental simulations of cultural transmission can be similarly useful in explicating the complexities of cultural evolution.

Acknowledgments. We thank Dan Glover for providing projectile-point images for the software program, David Geary for the use of his laboratory facilities, and Jelmer Eerkens and two anonymous reviewers for useful comments on the manuscript.

References Cited

- Arnold, Stevan J., Michael E. Pfrender, and Adam G. Jones
2001 The Adaptive Landscape as a Conceptual Bridge between Micro- and Macroevolution. *Genetica* 112–113:9–32.

- Barton, C. Michael, and Geoffrey A. Clark (editors)
1997 *Rediscovering Darwin: Evolutionary Theory in Archeological Explanation*. Archeological Papers No. 7, American Anthropological Association, Arlington, Virginia.
- Barton, Nicholas H., and Brian Charlesworth
1998 Why Sex and Recombination? *Science* 281:1986–1990.
- Bettinger, Robert L., and Martin A. Baumhoff
1982 The Numic Spread: Great Basin Cultures in Competition. *American Antiquity* 47:485–503.
- Bettinger, Robert L., and Jelmer Eerkens
1997 Evolutionary Implications of Metrical Variation in Great Basin Projectile Points. In *Rediscovering Darwin: Evolutionary Theory in Archeological Explanation*, edited by C. Michael Barton and Geoffrey A. Clark, pp. 177–191. Archeological Papers No. 7, American Anthropological Association, Arlington, Virginia.
- 1999 Point Typologies, Cultural Transmission, and the Spread of Bow-and-Arrow Technology in the Prehistoric Great Basin. *American Antiquity* 64:231–242.
- Boyd, Robert, and Peter J. Richerson
1985 *Culture and the Evolutionary Process*. University of Chicago Press, Chicago.
- 1989 Social Learning as an Adaptation. *Lectures on Mathematics in the Life Sciences* 20:1–26.
- 1992 How Microevolutionary Processes Give Rise to History. In *Evolution and History*, edited by Michael Nitecki, pp. 179–209. University of Chicago Press, Chicago.
- 1995 Why Does Culture Increase Human Adaptability? *Ethology and Sociobiology* 16:125–143.
- 2005 *The Origin and Evolution of Cultures*. Oxford University Press, Oxford.
- Camerer, Colin F., and Robin M. Hogarth
1999 The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk and Uncertainty* 19:7–42.
- Cavalli-Sforza, Luigi Luca, and Marcus W. Feldman
1981 *Cultural Transmission and Evolution*. Princeton University Press, Princeton, New Jersey.
- Chen, Peter C., and Paula M. Popovich
2002 *Correlation: Parametric and Nonparametric Measures*. Sage, Thousand Oaks, California.
- Cheshier, Joseph, and Robert L. Kelly
2006 Projectile Point Shape and Durability: The Effect of Thickness:Length. *American Antiquity* 71:353–363.
- Colegrave, Nick
2002 Sex Releases the Speed Limit on Evolution. *Nature* 420:664–666.
- Colegrave, Nick, and Angus Buckling
2005 Microbial Experiments on Adaptive Landscapes. *Bioessays* 27:1167–1173.
- Collard, Mark, Stephen Shennan, and Jamshid J. Tehrani
2006 Branching, Blending, and the Evolution of Cultural Similarities and Differences among Human Populations. *Evolution and Human Behavior* 27:169–184.
- Darwin, Charles
1859 *On the Origin of Species by Means of Natural Selection; or the Preservation of Favoured Races in the Struggle for Life*. Murray, London.
- Dobzhansky, Theodosius
1937 *Genetics and the Origin of Species*. Columbia University Press, New York.
- Dunnell, Robert C.
1980 Evolutionary Theory and Archaeology. *Advances in Archaeological Method and Theory* 3:35–99.
- Eerkens, Jelmer W., Robert L. Bettinger, and Richard McElreath

- 2005 Cultural Transmission, Phylogenetics, and the Archaeological Record. In *Mapping Our Ancestors: Phylogenetic Methods in Anthropology and Prehistory*, edited by Carl P. Lipo, Michael J. O'Brien, Mark Collard, and Stephen J. Shennan, pp. 169–183. Aldine, New York.
- Eerkens, Jelmer W., and Carl P. Lipo
2005 Cultural Transmission, Copying Errors, and the Generation of Variation in Material Culture and the Archaeological Record. *Journal of Anthropological Archaeology* 24:316–334.
- Elena, Santiago F., and Richard E. Lenski
2003 Evolution Experiments with Microorganisms: The Dynamics and Genetic Bases of Adaptation. *Nature Reviews Genetics* 4:457–469.
- Fisher, Ronald A.
1930 *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- Flenniken, J. Jeffrey, and Anan W. Raymond
1986 Morphological Projectile Point Typology: Replication Experimentation and Technological Analysis. *American Antiquity* 51:603–614.
- Fu, Wai-Tat, and Wayne D. Gray
2004 Resolving the Paradox of the Active User: Stable Sub-optimal Performance in Interactive Tasks. *Cognitive Science* 28:901–935.
2006 Suboptimal Tradeoffs in Information Seeking. *Cognitive Psychology* 52:195–242.
- Gurerk, Ozgur, Bernd Irlenbusch, and Bettina Rockenbach
2006 The Competitive Advantage of Sanctioning Institutions. *Science* 312:108–111.
- Henrich, Joseph
2004 Demography and Cultural Evolution: How Adaptive Cultural Processes Can Produce Maladaptive Losses: The Tasmanian Case. *American Antiquity* 69:197–214.
- Henrich, Joseph, and Robert Boyd
1998 The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. *Evolution and Human Behavior* 19:215–241.
- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis (editors)
2004 *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford University Press, New York.
- Henrich, Joseph, and Francisco J. Gil White
2001 The Evolution of Prestige: Freely Conferred Deference as a Mechanism for Enhancing the Benefits of Cultural Transmission. *Evolution and Human Behavior* 22:165–196.
- Henrich, Joseph, and Richard McElreath
2003 The Evolution of Cultural Evolution. *Evolutionary Anthropology* 12:123–135.
- Huxley, Julian S.
1942 *Evolution, the Modern Synthesis*. Allen and Unwin, London.
- Insko, Chester A., Robert Gilmore, Stephen Drenan, Andrew Lipsitz, Diane Moehle, and John W. Thibaut
1983 Trade Versus Expropriation in Open Groups: A Comparison of Two Types of Social Power. *Journal of Personality and Social Psychology* 44:977–999.
- Insko, Chester A., John W. Thibaut, Diane Moehle, Mark Wilson, William D. Diamond, Robert Gilmore, Marcus R. Solomon, and Andrew Lipsitz
1980 Social Evolution and the Emergence of Leadership. *Journal of Personality and Social Psychology* 39:431–448.
- Kameda, Tatsuya, and Daisuke Nakanishi
2002 Cost-Benefit Analysis of Social/Cultural Learning in a Nonstationary Uncertain Environment: An Evolutionary Simulation and an Experiment with Human Subjects. *Evolution and Human Behavior* 23:373–393.
- 2003 Does Social/Cultural Learning Increase Human Adaptability? Rogers's Question Revisited. *Evolution and Human Behavior* 24:242–260.
- Knecht, Heidi (editor)
1997 *Projectile Technology*. Plenum, New York.
- Lande, Russell
1986 The Dynamics of Peak Shifts and the Pattern of Morphological Evolution. *Paleobiology* 12:343–354.
- Lenski, Richard E., and Michael Travisano
1994 Dynamics of Adaptation and Diversification—a 10,000-Generation Experiment with Bacterial Populations. *Proceedings of the National Academy of Sciences* 91:6808–6814.
- Lipo, Carl P., Mark E. Madsen, Robert C. Dunnell, and Tim Hunt
1997 Population Structure, Cultural Transmission and Frequency Seriation. *Journal of Anthropological Archaeology* 16:301–333.
- Lipo, Carl P., Michael J. O'Brien, Mark Collard, and Stephen J. Shennan (editors)
2005 *Mapping Our Ancestors: Phylogenetic Approaches in Anthropology and Prehistory*. Aldine, New York.
- Lyman, R. Lee, and Michael J. O'Brien
1997 The Concept of Evolution in Early Twentieth Century Americanist Archaeology. In *Rediscovering Darwin: Evolutionary Theory in Archeological Explanation*, edited by C. Michael Barton and Geoffrey A. Clark, pp. 21–48. Archeological Papers No. 7, American Anthropological Association, Arlington, Virginia.
- 1998 The Goals of Evolutionary Archaeology: History and Explanation. *Current Anthropology* 39:615–652.
- Mayr, Ernst, and William Provine (editors)
1980 *The Evolutionary Synthesis*. Harvard University Press, Cambridge, Massachusetts.
- McElreath, Richard, Mark Lubell, Peter J. Richerson, Timothy M. Waring, William Baum, Edward Edsten, Charles Effereson, and Brian Paciotti
2005 Applying Evolutionary Models to the Laboratory Study of Social Learning. *Evolution and Human Behavior* 26:483–508.
- Mendel, Gregor
1866 Versuche Über Pflanzen-Hybriden. *Verhandlungen des Naturforschenden Vereines, Brünn* 4:3–47.
- Mesoudi, Alex
2007 Using the Methods of Experimental Social Psychology to Study Cultural Evolution. *Journal of Social, Evolutionary and Cultural Psychology*, 1:35–58.
2008 The Experimental Study of Cultural Transmission and Its Potential for Explaining Archaeological Data. In *Cultural Transmission and Archaeology: Issues and Case Studies*, edited by Michael J. O'Brien. The SAA Press, Washington, D.C., in press.
- Mesoudi, Alex, and Michael J. O'Brien
2008 The Cultural Transmission of Great Basin Projectile Point Technology II: An Agent-Based Computer Simulation. *American Antiquity*, in press.
- Mesoudi, Alex, Andrew Whiten, and Kevin N. Laland
2004 Is Human Cultural Evolution Darwinian? Evidence Reviewed from the Perspective of *The Origin of Species*. *Evolution* 58:1–11.
2006 Towards a Unified Science of Cultural Evolution. *Behavioral and Brain Sciences* 29:329–383.
- Morgan, Thomas H.
1932 *The Scientific Basis of Evolution*. Norton, New York.
- Neiman, Fraser D.

- 1995 Stylistic Variation in Evolutionary Perspective: Inferences from Decorative Diversity and Interassemblage Distance in Illinois Woodland Ceramic Assemblages. *American Antiquity* 60:7–36.
- O'Brien, Michael J. (editor)
1996 *Evolutionary Archaeology: Theory and Application*. University of Utah Press, Salt Lake City.
- O'Brien, Michael J., John Darwent, and R. Lee Lyman
2001 Cladistics Is Useful for Reconstructing Archaeological Phylogenies: Palaeoindian Points from the Southeastern United States. *Journal of Archaeological Science* 28:1115–1136.
- O'Brien, Michael J., and Thomas D. Holland
1995 Behavioral Archaeology and the Extended Phenotype. In *Expanding Archaeology*, edited by James M. Skibo, William H. Walker, and Axel E. Nielson, pp. 143–161. University of Utah Press, Salt Lake City.
- O'Brien, Michael J., and R. Lee Lyman
2000 *Applying Evolutionary Archaeology*. Kluwer Academic/Plenum, New York.
- 2002 Evolutionary Archeology: Current Status and Future Prospects. *Evolutionary Anthropology* 11:26–36.
- 2003a *Cladistics and Archaeology*. University of Utah Press, Salt Lake City.
- 2003b *Style, Function, Transmission*. University of Utah Press, Salt Lake City.
- Odell, George H., and Frank Cowan
1986 Experiments with Spears and Arrows on Animal Targets. *Journal of Field Archaeology* 13:195–212.
- Rice, William R.
2002 Experimental Tests of the Adaptive Significance of Sexual Recombination. *Nature Reviews Genetics* 3:241–251.
- Richerson, Peter J., and Robert Boyd
2005 *Not by Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press, Chicago.
- Schiffer, Michael B., and James M. Skibo
1987 Theory and Experiment in the Study of Technological Change. *Current Anthropology* 28:595–622.
- Schluter, Dolph, and Peter R. Grant
1984 Determinants of Morphological Patterns in Communities of Darwin's Finches. *American Naturalist* 123:175–196.
- Schluter, Dolph, Thomas D. Price, and Peter R. Grant
1985 Ecological Character Displacement in Darwin's Finches. *Science* 227:1056–1059.
- Schwartz, Theodore, and Margaret Mead
1961 Micro- and Macro-Cultural Models for Cultural Evolution. *Anthropological Linguistics* 3:1–7.
- Shennan, Stephen J.
2002 *Genes, Memes and Human History*. Thames and Hudson, London.
- Shennan, Stephen J., and James R. Wilkinson
2001 Ceramic Style Change and Neutral Evolution: A Case Study from Neolithic Europe. *American Antiquity* 66:577–593.
- Simpson, George G.
1944 *Tempo and Mode in Evolution*. Columbia University Press, New York.
- Steward, Julian H.
1938 *Basin-Plateau Aboriginal Sociopolitical Groups*. Smithsonian Institution Press, Washington, D.C.
- Wright, Sewall
1932 The Roles of Mutation, Inbreeding, Crossbreeding and Selection in Evolution. *Proceedings of the Sixth International Congress of Genetics* 1:356–366.

Notes

1. Although there have been attempts at a similar synthesis for human culture in the past (e.g., Schwartz and Mead 1961), these have been hindered by the flawed and distorted Spencerian theory of cultural "evolution" prominent within anthropology for much of the twentieth century (Lyman and O'Brien 1997).

2. As originally defined by Boyd and Richerson (1985:243), indirect bias can involve the selection of a model on the basis of any "indicator trait," of which hunting success is but one example (others might include prestige, similarity, age, health or sex: see Henrich and McElreath 2003). For brevity, and to maintain continuity with Bettinger and Eerkens (1999), we refer throughout simply to "indirect bias," although we mean "hunting-success-based indirect bias."

3. This argument is similar to Fisher's (1930) argument for the evolution of sexual reproduction in biological organisms. Fisher argued that sexual reproduction allows different beneficial mutations that have arisen in separate individuals to be brought together through recombination. In asexual organisms, in contrast, different beneficial mutations will not co-occur unless they independently evolve in the same individual lineage. Hence, both asexual reproduction (biology) and individual learning (culture) will drive individuals to locally optimal peaks in an adaptive landscape, whereas both sexual reproduction/recombination (biology) and indirectly biased horizontal cultural transmission (culture) allow individuals to jump to higher peaks in the adaptive landscape. Subsequent models and experiments have disagreed as to the validity of Fisher's (1930) hypothesis, and there is an ongoing debate regarding the adaptive advantages of sexual reproduction (Barton and Charlesworth 1998; Rice 2002). Generally, however, many of the problems with Fisher's hypothesis as applied to biological evolution may not apply to the case of cultural transmission, e.g., the two-fold "cost of sex" (cultural transmission does not produce/require two distinct "sexes"), the need for large numbers of beneficial mutations in the population (culture often has high mutation rates), and the fact that recombination is just as likely to break up combinations of beneficial mutations as bring them together (cultural transmission is not random, and people do not randomly copy only half of an individual's behavior).

4. Note that in Eerkens and Lipo's (2005) models the source of variation (copy errors in transmission) is different to the source of variation in our experiment (idiosyncratic individual learning in a multimodal fitness landscape). Nevertheless, indirectly biased cultural transmission has the same effect of reducing within-group variation irrespective of the source of variation.

5. The experimental design is most easily understood by viewing video clips of a participant engaging in each of the three phases. These videos can be viewed on the authors' websites, at either <http://amesoudi.googlepages.com/arrowhead-videos> or <http://cladistics.coas.missouri.edu/arrowhead-videos.html>.

6. We also ran an agent-based computer simulation of our experimental results, in which computer-generated agents played the same experiment as did our human participants

(Mesoudi and O'Brien 2008). To maximize the validity of the agent-based simulation, the behavior of the agents was specified using parameters from our experimental data. Hence the agents' individual learning was specified by the modal values of d and c exhibited by our human participants, and the agents' social learning was specified by the indirect bias exhibited by our human participants. After confirming that the simulated agents generated the same data as did our human participants, we then extended the simulation in ways that were impractical in the experiment, for example by increasing the group size beyond six, increasing the number of hunts beyond 30, changing the shape of the fitness landscape from multimodal to unimodal, and exploring the effects of alternative social learning strategies such as conformity or random copying. Subsequent footnotes highlight useful findings from this model; for more details see the forthcoming Mesoudi and O'Brien (2008), or contact the authors.

7. The agent-based simulation (Mesoudi and O'Brien 2008; see note 6) showed that the relative advantage of indirectly biased cultural transmission over individual learning also increased with group size. This is because with more people in a group, it is more likely that one of them will find themselves, perhaps by chance, at the globally optimal peak in the fitness landscape. Other group members can then copy this most-successful group member and all converge on the globally optimal peak. This provides another potential explanation for the difference between Californian and Nevadan points—a larger population size (or more accurately, greater population density) in Nevada could have made cultural transmission relatively more adaptive and hence more frequent, resulting in higher attribute correlations in Nevada compared to California, where low population density made cultural transmission relatively less adaptive and less common. There is evidence, however, for greater population density in California than in Nevada (Steward 1938), counting against this “increased Nevadan population density” hypothesis, so we favor the “more costly Nevadan individual learning” explanation given in the main text. (Note that these two explanations are not independent: a harsh environment could both increase the cost of individual learning and reduce population density.) This is a nice example of how experiments and models that draw on cultural transmission theory can generate competing hypotheses with clear predictions that can be tested with anthropological data.

8. The agent-based model (Mesoudi and O'Brien 2008; see note 6) showed that indirectly biased cultural transmission was more adaptive than a number of other social-

learning strategies (conformity, random copying, and copying the average), especially in larger groups (see note 7). The simulations also showed that the multimodal adaptive landscape assumption was key to this advantage. For example, conformity performed no better than individual learning in the simulations because there is no guarantee that the majority of individuals will be on the globally optimal peak. Agents engaging in conformist transmission are therefore just as likely to converge on a local optimum as a global optimum in the absence of information regarding the success of other individuals (unless individuals at the global optimum out-compete individuals at the local optima and become the majority, through cultural group selection). This contrasts with previous models that suggest that conformist transmission is adaptive under a wide range of conditions (Henrich and Boyd 1998), possibly because those models assume that individuals exhibit only one of two behaviors, one of which has a higher payoff, rather than a multimodal adaptive landscape with locally optimal fitness peaks. Indeed, our simulations showed that, in a unimodal adaptive landscape, conformity performed equally as well as indirect bias. However, the simulations also showed that, over large numbers of hunts in a unimodal adaptive landscape, neither conformity nor indirect bias performed better than pure individual learners, because in a unimodal adaptive landscape there are no locally optimal peaks for individual learners to get stuck on, and they will all eventually find their way to the top of the single globally optimal fitness peak. Generally, then, both our experiment and simulations show how the shape of the adaptive landscape can strongly affect cultural dynamics. The apparent paradoxes of stable suboptimal behavior frequently found by cognitive psychologists (e.g., Fu and Gray 2004, 2006) might also be resolved if participants in those experiments were allowed to engage in indirectly biased cultural transmission, and jump to a global optimum.

9. Indeed, this is very similar to the biological evolution of beak size and shape in Darwin's finches, in which there is a “trade-off between feeding rates on small soft seeds and large hard seeds” (Schluter et al. 1985:1057). Schluter and Grant (1984) showed that such trade-offs result in a multimodal adaptive landscape, the shape of which has significantly influenced finch-beak evolution.

*Received February 1, 2007; Revised March 14, 2007;
Accepted March 28, 2007.*