UNDERSTANDING CULTURAL EVOLUTIONARY MODELS

A REPLY TO READ'S CRITIQUE

Joseph Henrich

Joseph Henrich, Department of Anthropology, Emory University, 1557 Dickey Drive, Atlanta, GA 30322 (jhenric@emory.edu)

This reply to Read's (2005) critique of my paper (Henrich 2004) is divided into three parts. Part I clarifies Read's misinterpretations and mischaracterizations of both Dual Inheritance Theory in general and my model specifically. Part II addresses several problems in Read's empirical analyses of forager toolkits, and presents an alternative analysis. Part III tackles some common misunderstandings about the relationship between cost-benefit models (such as Read's) and cultural evolutionary modeling approaches, as well as highlighting some concerns with Read's efforts. In writing this, I have tried to introduce the reader to the issues in debate, but to fully understand this reply, one should read both my paper and Read's critique.

MISCHARACTERIZATIONS AND MISINTERPRETATIONS

Part I addresses four aspects of Read critique: (1) Read's confusion about the breadth and applicability of cultural transmission models, (2) the misdirected nature of Read's criticisms of certain specific modeling assumptions, (3) the lack of important differences, contra Read, created by partitioning the "knowledge" and "technique" components of skill, and (4) the misunderstandings inherent in Read's claims that my model requires unrealistic parameters.

The Breadth and Applicability of Evolutionary Models of Cultural Transmission

Read premises much of his criticism, and the foundations of his later graphical discussion, on the idea that my model—and Dual Inheritance Theory more generally—applies only to the "imitation" of "motor skills," and not to the cultural transmission of knowledge.[1] This is a deep misunderstanding. To set the record straight, the logic of my model applies to any cultural domain influenced by (1) success and/or prestige-biased transmission and (2) imperfect learning. Since we know empirically from the vast literature in Psychology that both "knowledge" and "motor skills" are influenced by prestige and success-biased transmission, and that these processes result in imperfect transmission, my model is set on a solid empirical foundation (for reviews of the literature see Henrich and Gil-White 2001; Henrich and Henrich in press: Chapter 2). Since my paper was aimed at archaeologists and sought to use the Tasmania case as an example, I favored the word "skill" throughout, and wrote with technological products in the foreground. "Skill," as I used it, included both knowledge and technique, and assumes learners observe both verbal (e.g. speaking) and non-verbal behavior in their "cultural models" (i.e., the individuals they are learning from) in acquiring their final mental representations.[2] In the case of technology, these skills often manifest themselves in the material record. However, I emphasize that nothing about the model confines it to the domain of technology: one could, with the right dataset, analyze the effects of demographics on cultural domains related to folkbiological knowledge.

Early in his paper, Read leads (e.g., p.5) the reader to believe that by "skills" I mean "motor skills"—which I did not mean (see below). While Read starts sliding the word "motor" in front of "skill" in describing my model, the word "motor" never appears in my paper. Read's misunderstanding is particularly odd since later in his own paper, in discussing his own model, he uses the word "skill" in the same way I did. Read writes, "with the exception of tasks requiring a high level of skill (both in terms of motor development and knowledge about effective task performance)…" [2005: 39].

Read also mischaracterizes Dual Inheritance Theory in general, which is much harder to do given that there are multiple books and dozens of papers specifically on the topic. Dual Inheritance theory's cultural evolutionary models apply to a wide range of cultural domains, including culturally-transmitted skills, beliefs, knowledge, preferences, practices, techniques, and strategies. These models are general and can include the influence of verbal behaviors (e.g., words and sentences) and/or non-verbal behaviors on cultural transmission. As reviewed in Boyd and Richerson (1985: Chapter 3), a great deal of research on social learning suggests that learners use both the observation of non-verbal actions and their model's words to facilitate cultural learning (Bandura 1977). Since both verbal and non-verbal behaviors go into the transmission process, it would be a mistake to exclude either. Furthermore, since the transmission process *integrates* both verbal and non-verbal behavior, it is not clear that partitioning these into separate channels (as Read suggests, without reference to empirical psychology) is necessary or sensible. A learner's knowledge is not merely a recording of what people tell him. Learners are selective in who they listen to, and acquired knowledge is substantially edited based on the non-verbal behavior of speakers. For example, if an expert tells a learner that a particular type of wood is best for making arrows, but the expert himself uses a different type of wood, learners will likely defer to the non-verbal action as a source of knowledge. This effect has been demonstrated in laboratory studies of social learning (reviewed in Henrich and Henrich in press: Chapter 2). Note also that Read's intuitions about knowledge transmission as straightforward "communication" does not fit with what is known empirically about cultural learning in small-scale societies (Fiske 1998; Lancy 1996).

To confront these confusions, I highlight material from three sources: (1) my paper, (2) the main theoretical and empirical source cited in my paper, and (3) the foundational work in Dual Inheritance Theory cited by Read. Let's begin with two quotations from my paper. First, from the section in my paper that describes the construction of the formal model, and specifically the evolving cultural variable, z_i, I write:

> Transmittable z_i skills might involve such things as net-manufacturing *preferences* (weaving practices, *preferences* for certain fibers), spear-throwing techniques, fishhook

> material selection, canoe-building techniques, bone-tool craft, and *medicinal plant knowledge* [Henrich 2004:200; emphasis added].

Clearly, I did not confine the model to "motor skill" or exclude knowledge, preferences, etc.

Later, while arguing that individual learning would not, on average, improve an individual's z_i measure, I write:

> …the ethnographic record … suggests that these circumstances are unlikely to apply to most human situations because of the amount of *culturally learned know-how* involved in human skills related to making and using such things as blowguns, bows, arrows, bowls, craft tools, spears, fishing nets, canoes, kayaks, etc., or in practices and *knowledge* related to such things as tracking, using medicinal plants, and processing foraged foods [Henrich 2004:202; emphasis added].

Given these quotations, and many others in my paper, it is hard to see how Read concluded that the model was about "motor skills," and did not include "knowledge."

Since the theoretical and empirical point of departure for my paper was Henrich and Gil-White (2001), it makes sense to examine this paper. The first line of the abstract refers to "information acquired via cultural transmission." Later in this paper, we write:

> …our explanation focuses on particular forms of *direct* social learning, which we collectively term *infocopying*. This category encompasses all forms of acquiring information directly from another, and includes, but is not limited to, "true imitation" (acquiring the details of motor patterns via direct observation; see Tomasello 1994) and "goal emulation" (inferring behavioral goals via direct observation). Infocopying excludes indirect social learning processes, such as "social facilitation" or "local enhancement," where learners have a higher probability of reinventing something due to

> close proximity to a competent performer and the materials involved. Infocopiers may also unconsciously acquire mannerisms, consciously acquire verbal knowledge and arguments, and consciously or unconsciously imitate action patterns…[Henrich and Gil-White 2001:172; emphasis original].

Again, it is hard to see how Read got the idea that the model, and my work more generally, was about replicating motor skills.

Read also references Boyd and Richerson (1985) and suggests that somehow their models of cultural transmission do not include knowledge transmission, or the influence of verbal behaviors (like talking) on transmission. This claim is contradicted in so many places in their book that it is hard to know where to begin. Two quotations will suffice. First, their discussion of culture emphasizes this: "Culture is information capable of affecting individuals' phenotypes which they acquire from other conspecifics by teaching or imitation (1985: 33)… The essential feature of culture is social learning, the nongenetic transfer of patterns of skill, thought, and feeling from individual to individual in a population or society" (1985:34).

Now, some confusion might result from the manner in which both Boyd and Richerson, and I, sometimes use the word "imitation" (in its various forms) as a handy gloss for a variety of complex forms of cultural transmission. Those not familiar with the technical usage of the word "imitation" in the psychology literature (Bandura 1977; Tomasello 1999) may tend to unconsciously reinterpret our more technical application of the term based on their own intuitions from common English usage. However, our use of the gloss "imitation" is always accompanied by technical definitions and illustrative examples, meant to clarify our meaning, and suppress this reinterpretation.

Criticisms specific to my model and assumptions[3]

Read makes quite a big deal about my modeling assumption that permits learners to locate and learn from the most successful individual in the overall population. While this is obviously an unrealistic assumption, that is *not* the standard that one should use in judging a modeling assumption! Every useful model is loaded with unrealistic assumptions—see below. A better question to ask is how does the assumption influences the model's results, and any claims arising from those results. In my case, the assumption is *highly conservative*, and works against the larger case made in my paper. My claim was that errors and noise in the transmission of complex skills have smaller effects in retarding adaptive cultural evolution in larger, more interconnected, populations than in smaller, less interconnected, populations. If the overall population is partitioned into smaller subpopulations (such that learners could only access the locally most skilled) connected by migration, the effective size of the interconnected pool of social learners (N in my model) would decrease, thereby *magnifying* the deteriorating effects of population size and interconnectedness that my paper highlights. Thus, I made a tactical assumption that simplified the mathematics and *maximized the difficulty* of showing what I showed. Releasing this assumption both makes the model more realistic and makes the process *more likely* to occur. This is why I wrote, in reference to the assumption in question, "In studying the conditions for maladaptive deteriorations in skill, this is a highly conservative assumption that favors cumulative cultural adaptation" [Henrich 2004:214]. The proper approach to addressing a concern about this assumption would have been to modify the model by releasing the assumption in some more realistic manner, and then demonstrating that one gets a qualitatively different result. Read did not do this.

Those unfamiliar with formal evolutionary modeling might not understand *the principle of optimal simplicity*. The idea is that a proper formal evolutionary model aims to capture the governing dynamics of a process, while at the same time stripping the process to the bone. Unrealistic assumptions are rampant in *every* important model. Analogous to Read's criticism of my efforts, models from evolutionary biology typically assume random mating within large (infinite!) populations. That is, they assume every animal is equally likely to mate with every other animal (e.g., wolves in North America). Read presumably would criticize all of these models for a lack of realism, despite the fact that in many applications this assumption is harmless.[4]

With respect to the Tasmanian case in particular, however, the applicability of my model does require that either culture (ideas, practices, skills or knowledge) or people move among the smaller subpopulations scattered around the island. If, for example, Paleolithic Tasmanian bands were not linked through social networks through which cultural information could flow across the Bassian peninsula to Victoria, then cutting off Tasmania from mainland Australia would not have influenced the size of the interconnected pool of social learners—and thus my model could not explain the losses.

There are several reasons to believe that at least certain kinds of culturally-transmitted information would have diffused widely through Tasmania.[5] First, evidence from other foraging people shows that skills, technology, and know-how diffuse over substantial geographic areas, and across ethno-linguistic boundaries (see Collard, et al. in press and references therein). Thus, to make the case that local Tasmanian groups were isolated from one another (as Read does), one would need to explain why Tasmanians were so different from other foragers or horticulturalists. Second, we know from the ethno-historical record that hunting with dogs spread rapidly across

Tasmania, so at least with respect to this hunting practice, subpopulations were well-interconnected (Jones 1977b:46). Third, Jones (1995) summarizes a variety of evidence for inter-island trade. For example, he writes (1995:426-427), "About 2.5kyr ago the exotic stone material includes a fine-quality spongolite chert from a quarry near the west coast at Rebecca Creek, 80km away (31). Artifacts from this source were widely found in middens dated to the past two thousand years along the western and northern coasts of Tasmania." Similarly, Jones (1995:435) explains that, "The geographic distribution of Darwin glass over an area of 15,000 km^2 gives an indication of the scale of movement of people who at some stage of their lives had recourse to the same economic source, whether by trade or direct access."

Read's claim that my model required that every learner imitate the most skilled individual in the population misses an entire subsection of my paper that incorporated an arbitrary amount of vertical cultural transmission, along with prestige or success-biased cultural transmission. This modeling effort can be construed in one of two ways. In the first approach, all learners initially acquire their skill (including knowledge, techniques, goals, etc.) from one of their parents (assume it's their same-sex parent). Then, later, a proportion p of the population learns from a highly successful or prestigious model. As explained above, I make the tactically-appropriate assumption that they choose the most successful/skillful individual in the population form whom to learn. Alternatively, this setup could equally well be interpreted as assuming that a proportion 1- p of the population learn from their same-sex parent, and a proportion p learn from the most successful individual in the population. Either way, the derivation shows that the *basic qualitative findings do not change* (as long as $p > 0$), even if most people are *not* learning from the most successful individuals, and are instead learning from their parents.

Read's appendix claims to generalize my derivation to any distribution of skill. The problem with this is that *my* appendix already contains this finding: In Appendix A, just prior to assuming a Gumbel distribution, equation (A2) is already at about the same place Read ends up. However, my derivation provides a greater understanding of the evolutionary issues, since it emerges from the Price Equation (the general expression of evolutionary change for any system). This connection allows one to readily see what would happen if the "copy the most successful individual in the population" assumption is relaxed so that learners are only somewhat inclined towards more successful individuals. Moreover, Read appears to have missed the significance of the Gumbel distribution. The Gumbel distribution is the distribution one robustly gets if the highest value is repeatedly taken from samples of size N. Only by choosing an appropriate distributional form can we press forward and derive testable predictions. Read's appendix derivation does not get us any scientific traction: It makes no non-intuitive predictions and does not yield any rough-and-ready equations that relate demography to evolutionary dynamics.

Partitioning knowledge and technique does not matter

While, as I explain below, Read's efforts at partitioning "knowledge" and "motor skill" do not get us very far, the material I presented in Appendix D of my paper does allow us to partition z (e.g., skill, success, return rate) into various components, such as "knowledge" and "technique." Here I use "technique" instead of Read's "motor skill." Suppose, for illustrative purposes, z presents hunting returns, a phenotypic measure used by individuals in selecting who to learn from; suppose y is an underlying measure of knowledge that can be acquired (approximated) by watching, listening, and inferring from a chosen model's verbal and non-verbal behaviors, and ϕ captures an individual's technique, which is also inferred by learners from observation and possibly instruction. Using the λ values in equation (1) below we can express the relative contributions of individual *i*'s knowledge and technique to his hunting returns (z_i).[6]

$$z_i = \mu + \lambda_1 y_i + \lambda_2 \phi_i + \varepsilon_i \tag{1}$$

Here, the λ's give the relative contribution of an individual's knowledge, y, and technique, ϕ, to his observed returns, z_i. The parameter ε_i is an uncorrelated random error, and μ is a constant.

Now, following the derivations in the Appendix D of my paper, we define f as an individual's relative cultural fitness, which determines their likelihood of being selected as a model by a learner. An individual's relative cultural fitness depends on their hunting returns (z_i), other factors (summarized in x_i), and uncorrelated random error (e_i):

$$f_i = m + \rho_1 z_i + \rho_2 x_i + e_i \tag{2}$$

Here, ρ_1 is the partial regression coefficient of f on z: the relative contribution of an individual's hunting returns to his relative cultural fitness. Putting this into the basic Price formulation for $\Delta\bar{y}$ and $\Delta\bar{\phi}$ (that is, the change in knowledge and technique per time step) yields:

$$\Delta\bar{y} = \mathrm{Cov}\left(f, y\right) + E\left(f\Delta y\right) = \rho_1 Cov\left(z, y\right) + E\left(f\Delta y\right)$$

$$\Delta\bar{\phi} = \mathrm{Cov}\left(f, \phi\right) + E\left(f\Delta \phi\right) = \rho_1 Cov\left(z, \phi\right) + E\left(f\Delta \phi\right)$$

Including the causal relation of y_i and ϕ_i with z_i, we arrive at the following:

$$\Delta\bar{y} = \lambda_1 \rho_1 \mathrm{Var}(y) + E(f\Delta y);\ \Delta\bar{\phi} = \lambda_2 \rho_1 \mathrm{Var}(\phi) + E(f\Delta \phi)$$

This brings us right back to where my original paper started, with the exception of the constants (ρ and λ). That is, partitioning knowledge and technique does not change the qualitative insights. This finding shows that either the transmission of knowledge or technique could prove to be the limiting component in the evolution of skills (and any associated technologies). Which one on these sets the limit would depend on the relative strengths of selective forces (the first term on the right-hand side in the above equations) and the effects of low transmission fidelity (the second term on the right-hand side).

"Unrealistic Parameters"

Read's discussion of what he calls the "unrealistic parameters" for learning in my model is both misleading and misdirected. First, perhaps the oddest aspect of this claim is that it comes on the heels of his criticism about the assumption that learners can imitate the most skilled individual in the population. If we weaken learners' ability to find the most skill individuals in the population, then the ratio of negatively biased error (α) to the spread in the error distribution (β) can be substantially lower, and still create technological losses and related demographic effects. The particular values that Read intuitively finds too high are an artifact of the conservative assumption I used in solving the model. It is important to realize that formal models are used for different purposes. In my case, as is often the case with such models, the goal was to illustrate the logic of an argument and derive some non-intuitive qualitative predictions. If I had wanted to estimate the individual learning and error parameters (α and β), I would have included at least one parameter for the cultural learning part of the process.

Next, while weakening cultural learners' abilities in this way addresses Read's intuitive concerns about α/β, I must note that I do not find the required values unrealistic at all, especially for complex skills. All that is required is that the modal loss in skill be four or more times greater than the statistical spread around that mode (see Figure 1 in my paper). Imagine a group of 1000 novice archers trained by a grandmaster, who himself scores 150 on the required archery target examination. After training, most of these newly fledged archers score between 20 and 40 on the examination (before training most cannot hit the target at all, so this is a big improvement). Here

$z_h = 150$, $\alpha = 120$, $\beta = 20$, and $\alpha/\beta = 6$. Does this example seem implausible? What would happen if, for each incoming class of novices, the previous grandmaster was replaced by the best student (who now gets the title of grandmaster) from the most recent class of graduates? What would happen if the class size was reduced from 1000 to 10?

EMPIRICAL EFFORTS: FORAGER TOOLKIT COMPLEXITY

Combining the available data on hunter-gatherer demographics with data on the complexity of food-getting technology is a potentially important avenue for examining hypotheses about the evolution of technology—interested readers should see Collard et. al. (2005). However, I have four problems with Read's foray into this. These are: (1) the applicability of the particular demographic measures Read uses to the issue at hand, (2) the integration of these demographic data with the tool complexity data by matching timeless ethnolinguistic group labels, (3) Read's operationalization of my model with regard to an appropriate measure of toolkit complexity, and (4) the fact that Read arbitrarily altered key pieces of data without justification. Below, I will take each of these in turn, and then show how we can use these data to make some headway.

Read uses population densities gleaned from Binford (2001) as predictors of measures of toolkit complexity from Oswalt (1976), in his effort to test my model. There are two problems here, one dealing with the general appropriateness of "population size" and "population density" vis-à-vis my model, and a second focusing these variables as Binford defined them. As to the first, testing my model requires a measure of my N, which is the size of the pool of interacting social learners. This variable integrates both population size and its degree of interconnectedness. Population density obviously is not the right measure since, for example, small islands can have high densities with only small numbers of learners. Or, alternatively, highly nomadic groups might remain well-connected across a large population for many domains of cultural transmission despite living at low densities. For example, widely scattered groups of Kalahari foragers share many elements of technology and know-how that are—we know from ethnography—culturally-learned, despite living at low density. Population size, though not used by Read, is another variable provided by Binford, and has been used by Collard et. al. (2005) to explore toolkit complexity. This brings us to the second, closely related, point. Binford (sensibly) defines "population size" and "area" in his data according to the group in the ethnography from which he draws his data. This tells us nothing of the relevant population of interconnected social learners. Binford's Tasmania data, for example, comes from only two groups, with the overall population of Binford's Tasmania samples accounting for only about one-quarter of the overall Tasmanian population. That is, one cannot calculate the population size or density of Tasmania from Binford's data—so Read criticizes me for assuming the population of Tasmania is interconnected, but then when he attempts to test my model he does *not* use island-wide data.[7]

Second, to assemble his dataset Read had to match the demographic data from Binford and Torrence with the tool data from Oswalt using group labels. This risks matching-up data taken at very different places and times. As a spot check on this, I went to the ethnographic sources that Oswalt used for the Andamese to obtain his measure of food-getting toolkit complexity, pulled the population and densities information from this same source, and compared this to the population and density information found in Read. Radcliffe-Brown (1964: 16-17) estimates population densities ranging from 106 persons per 100 km^2 (North Andaman) to 67.5 (Little Andaman and North Sentinel), and provides an overall average density of 87. Read uses 42.3 for Andamese density, a value half the size of Radcliffe-Brown's measures. This difference could result from population data taken at different times in Andamese history, different methods by different ethnographers, or that fact that while Binford's values (and

Read's) are for specific local groups (based on ethnographies, not meant to capture overall regional values), Oswalt's tool measures often aggregate technology over much larger areas.[8]

The third problem with Read's empirical effort emerges from his attempt to operationalize my model. As explained in my paper, the effect of a reduced pool of interacting social learners will not hit all technologies (or knowledge) equally. Instead, it will preferentially target the most complex skills (i.e., tools that are hard to learn to make, and easy to screw up). Simple technologies won't be affected. Thus, measures such as the number of *subsistents* (number of different food-getting technologies), the total number of technounits (parts making up subsistents) across all subsistents, or the average number of technounits per subsistent all miss this differential targeting. In lieu of these, a more-appropriate measure is simply to sum up the technounits in those subsistents with the most technounits in each category, across Oswalt's four categories of food-getting technologies: instruments, weapons, tended facilities, and untended facilities (call this variable MXT). MXT allows us to analyze and compare the most complex technologies across our sample of societies. For example, in Tasmania the most complex instrument has 1 technounit (tu), the most complex weapon has 1 tu, the most complex untended facility has 4 tu, and the most complex tended facility has 1 tu. This gives Tasmania a MXT score of 7. Or, among the Yahgan, the most complex instrument has 6 tu, the most complex weapon has 6 tu, the most complex untended facility has 6 tu, and the most complex tended facility has 5 tu, giving MXT = 23.

Finally, before making use of this new dependent measure, I must highlight the fact that Read changed or selectively ignored the data he took from Binford. Since the Tasmanians are the group in question, the most problematic change was to raise the effective temperature (ET) from Binford's value 12.62 to Read's 19.12. As a climatic measure, ET integrates the annual distribution of solar radiation with its intensity. Mysteriously, Read decided that the Tasmanian ET value from Binford, *and only the Tasmanian value*, did not capture what ET is supposed to measure. To "fix" this, Read regressed ET on another of Binford's climatic variables, TEMP, to get a linear relationship. He then used this regression to recalculate the Tasmania ET value. There are two major problems with this. First, if Read thinks TEMP is a better measure, then use TEMP instead of ET in the toolkit complexity regression (or create and justify a new climatic variable; Binford provides several climate variables); but one should not arbitrarily change one unpleasant data point, especially when it is the key data point at issue. Any adjustments, besides requiring a real justification, should be uniformly applied to *all the data points*. In fact, if we apply the "correction" that Read applied only to the Tasmanian ET value (found in footnote 2 of Read's Table 1) to all the groups, we find that the Andamese in fact *differ by more* than the Tasmanians (but they were not "corrected"!), and other groups differ by almost as much. Why weren't they "corrected"? Second, putting aside the inappropriate and selective nature of this modification, Read should not have used a linear regression to create his "correction." Binford's Figure 4.05 clearly shows a sharply non-linear relationship between ET and TEMP. Third, Read's "correction" to ET forces us to re-classify Tasmania as "tropic" island (see Binford's Table 4.02). As they are buffeted by the roaring forties, I bet the inhabitants of Tasmania will be surprised to hear that they live in tropics. All of Read's subsequent comparative discussion of Tasmania, vis-à-vis the Fuegian groups for example, is afflicted with this "correction."

Read also decided to change the population density measures for the Chenchu and the Owens Valley Paiute. Read writes, without further justification, "Density given in Binford (2001) is too large." To "fix" this, Read regresses density on primary biomass, obtains a linear equation, and alters the densities accordingly. Read apparently has no tolerance for data that deviates from univariate linear ecological relationships. Read also decided to drop the Andamese out of his regression of toolkit complexity on ET because it “appears to have more technounits per tool type than would be expected [based on ET] due to the importance of meat in the diet…” [Read 2005:11]. Of course, as we will see below, Collard et. al. (2005) has already explored this and shown that the contribution of terrestrial animals (or aquatic animals) to the diet does not predict the toolkit variables that Read tested. It is unclear why Read did not run multivariate regressions, as Collard et. al. did.

Before analyzing MXT, a more appropriate measure of technological complexity vis-à-vis my model, the reader should note that no where did I claim that environment and ecology were not relevant to technology or technological change. To the contrary, the heart of cultural learning mechanisms like prestige-biased transmission is their ability to adapt to novel and uncertain environments by tapping the accumulated knowledge/skills, individual experiences, and innovations of others. Thus, our analysis begins by first explaining as much of the variability in MXT as we can using environment and ecology. To accomplish, I use Collard et. al.’s data plus the Yahgan. I prefer Collard’s data to Read’s, as the former has been synthesized and cross-checked across a variety of sources, including but not limited to Binford and Oswalt, and takes a multivariate approach. Collard et. al. also did not apply Read’s “corrections.” I added the Yahgan using data from Binford and Oswalt because they were specifically referred to by Read.

With these data, I regressed MXT on (1) effective temperature (ET), (2) above ground productivity, (3) number of residential moves per year, (4) distance traveled annually during residential moves, (5) percentage contribution of terrestrial animals to the diet, and (6) the percentage contribution of aquatic animals to the diet.[9] Following Collard et. al., I used a step-wise backward regression, but any approach will yield the same result (e.g., step-wise forward):

the only variable that explains any significant proportion of the variation in MXT is ET. As a lone predictor, ET captures 50% of the variation in MXT, with a standardized beta coefficient of -0.71 ($p < 0.001$).

Having removed as much of the influence of environment as we can, we can ask how each group's predicted MXT values compare to their actual MXT values. This residual difference is plotted for each group on Figure 1. The first thing to notice is that Tasmania remains dramatically below where ET would predict. In fact, Tasmania is the *only* data point that standard regression rules-of-thumb flag as an outlier. Tasmania's MXT value is 7, the ET regression predicts that Tasmanians ought to have an MXT value of 27.4--a factor of four difference. [Figure 1 about here]

While Tasmanians have, by far, the least complex food-getting tools, *especially* once their environment has been controlled for, a closer look at the other deviant groups is informative. On the low side, the Yahgan, who are predicted to have an MXT of 35 but show an actual value of 23, are the "canoe people" who inhabit the islands making up the archipelago of Tierra del Fuego. The rugged geography is such that they are only tenuously connected to the social networks of South America, and technological know-how would have to transmit through the land-dwelling Ona. Thus, their lower-than-expected MXT value could plausibly be a consequence of demographic factors.

The Tiwi have a MXT value of 6 and a predicted value of 14.6. Like the Tasmanians, the Tiwi are another island dwelling population, inhabiting Bathurst and Melville islands, some 30 miles off the coast of Northern Australia. According to Oswalt (1976: 163), the Tiwi believed mainland Australia to be the land of the dead, and likely had very little contact (also see Hart and Pilling 1960). Thus, rather than providing an exceptional case as Read suggests, the Tiwi possess a simple technology for *precisely the same reason* as the Tasmanians: isolation from the continental networks of cultural transmission.

Putting a finer edge on this point, the Groote-eylandt live on another island off the coast of Northern Australia, have the exact same ET values as the Tiwi, but are just a touch below their expected MXT value. Unlike either the isolated Tiwi or Tasmanians, this group—with a shorter ocean crossing—maintained contact with the mainland (Jones 1976), and thus did not suffer technologically.

Finally, while Read ignores them, the Andaman islanders may provide an important test case. These island dwellers show an actual MXT well above their predicted MXT, based on their ET value. Now, this could be a problem for my theory, unless these groups have had sufficient contact with either Indonesia or mainland Asia. A potentially important difference here is that the Andamese, rather than being off the coast of Australia (a continent of entirely foragers before Europeans arrived), lived 80 miles from the coast of Asia, home of the most technologically and politically complex societies on earth. Radcliffe-Brown (1964: 6-7) reports that the Andamese are mentioned in both Chinese and Japanese writings in the first millennium A.D, by Marco Polo (who heard about them second hand), and by two Arab travelers in 871 A.D. While not much to go on, these hints suggest a potentially important contrast with the Tasmanian and Tiwi cases.

In my paper, I explained that the real challenge for those who strictly adhere to ecological and economic models is to explain why the Tasmanians were so different--in terms of technological complexity—from their aboriginal cousins 150 miles across the Bass strait in Victoria. Read reaffirms my challenge when he writes:

> As can be seen in Table 2, these two regions [Victoria and Tasmania] are comparable with regard to ET and TEMP, and have similar *predicted* numbers of tools

> and technounits/tools. That Australian groups in Victoria had a more complex tool assemblages than was the case for the Tasmanians may thus [thus?] be explicable through the difference in climatic conditions between the two regions, taking into account the manner in which differences in climatic conditions affected the seasonality and spatial distribution of fauna and flora [Read 2005:14, emphasis added].

Read says nothing more on the comparative question. However, between Collard et. al.'s analysis, and my work above, we tried measures that take much of this into account, and show no explanatory effects. If Read has additional climatic or ecological variables, he should explore them (not vaguely suggest that some must exist). I recommend that while we are searching for micro-climatic variations between Victoria and Tasmania that might cause such substantial differences in technological complexity, we should continue to consider the fact that aboriginal Victorians were connected via social networks to a continent of 19,913,290 km^2, while aboriginal Tasmanians remained isolated on their island of 68,332 km^2 for 10,000 years.

EVOLUTIONARY MODELS VS. COST-BENEFIT GRAPHS

Read's alternative approach proposes that skills require practice, practice is costly in terms of time, and time is constrained. His graphical discussion suggests that a learner, balancing some measure of efficiency against costly practice, should under some circumstances prefer less efficiency (in skill or technology) because obtaining greater skill would require too much investment in learning. Read's discussion of this idea, vis-à-vis my model, requires numerous comments. Also, I should point the reader to two full-fledged formal economic models of technological change that go well beyond Read's graphical discussion (see Bettinger, et al. forthcoming; Ugan, et al. 2003).

Read's contrast of his economic explanation with my cultural evolutionary model reveals a common misunderstanding about the relationship between economic and cultural evolutionary models. To see this, let's begin by considering Read's idea from the perspective of a Tasmanian learner. Suppose there are two hypothetical hunting skills in Tasmania, one involves making and learning to throw clubs (Skill I) and the other involves making and learning to throw boomerangs (Skill II). Suppose not a single person currently performs Skill II on the island. Given this, how does our learner know the *costs* of learning to make and throw a boomerang? Even if there were a few boomerang-throwers on the island, all our learner might see would be lots of dead wallabies, but not the costs of learning the skill. To make the calculations required by Read's approach, every learner would have to begin the process of learning Skill II, and then pull out if need be, during the learning process. Otherwise, there is no way a learner could know that it is too costly to learn to make and use a boomerang. Neither the ethnographic nor the archaeological records reveals evidence that most people gather cost/benefit information about technology and craft production via this individual experimentation process (Fiske 1998; Lancy 1996). There are not, for example, lots of unfinished or poorly made bone tools in the later archeological record of Tasmania (Jones 1977a). There is simply no way that learners could know that bone tools were too costly to waste time investing in (supposing that they were). Now Read might reply that he did not mean to suggest that each learner individually experimented with each possible technology and skill to assess its costs and benefits. However, if he does think that much of this information is transmitted (in some fashion) among individuals, then he is forced back into the world of cultural evolutionary (learning) models. In fact, most economic models will transform into learning, and usually into cultural learning, models once the acquisition of information (about costs and benefits, or anything else) is not considered totally free.

Cultural evolutionary models, building on substantial empirical evidence, propose that learners often use cues that indirectly or implicitly contain information about costs and benefits.

These models provide psychologically plausible means by which individuals can adapt dynamically to the costs and benefits of technologies, physical environments, and their social worlds. For example, if an individual learns both how to invest his time, and some of the knowledge and techniques associated with that investment, from individuals who are highly successful, then he will be able to avoid investing in pursuits that are too costly (those investing in “too costly” pursuits won’t achieve great success). In the case of our Tasmanian learner, if Skill II is in fact too costly and leads to too little investment in other areas (e.g., tracking or social relationships), then those who invest in Skill II will not be imitated, and Skill II will not spread. Now, however, our individuals need not engage in trial and error learning to figure out that they should not invest in Skill II. Most individuals will never even consider learning Skill II, or need to. Approaching the problem of learning about costs and benefits in this indirect fashion means that the effects described in my model can apply, as well as other evolutionary forces like drift (Shennan 2001).

The final problem with Read’s model, and similar static economic approaches, is that there’s no evolution. The model has no endogenous dynamics. If costs and benefits change, does the entire population shift instantaneously to the new optimal technology? If cost and benefits, or climatic variables for that matter, do not change, is culture (including technology) static? To see this more sharply, consider that my model connects demographics to rates of cumulative cultural evolution. I highlighted the possibility that a drop in the size of the pool of interacting social learners could initiate a process of cultural loss that could include adaptive knowledge, practices, skill, etc. But, on the flip side, the model also predicts different rates of adaptive cumulative cultural evolution, depending on the size and interconnectedness of a population. It can also address why technology becomes increasingly adaptive and complex over broad time spans, without any direct connection to climate (McBrearty and Brooks 2000). It is unclear how Read's model addresses cumulative cultural evolution, possibly the hallmark of our species (Boyd and Richerson 1996).

References

Bandura, A.

1977 *Social learning theory*. Prentice Hall, Englewood Cliffs, N.J.

Bettinger, R. L., B. Winterhalder and R. McElreath

Forthcoming A simple model of technological intensification. *Journal of Archaeological Science*.

Binford, L. R.

2001 *Constructing Frames of Reference*. University of California, Berkeley.

Boyd, R. and P. Richerson

1996 Why Culture is Common, but Cultural Evolution is Rare. *Proceedings of the British Academy* 88:77-93.

Boyd, R. and P. J. Richerson

1985 *Culture and the Evolutionary Process*. University of Chicago Press, Chicago, IL.

Collard, M., M. Kemery and S. Banks

2005 Causes of Toolkit Variation Among Hunter-Gatherers: A Test of Four Competing Hypotheses. *Journal of Canadian Archaeology* 29:1-19.

Collard, M., S. J. Shennan and J. J. Tehrani

in press Branching, blending, and the evolution of cultural similarities and differences among human populations. *Evolution and Human Behavior*.

Fiske, A. P.

1998 Learning Culture The Way Informants Do: Observing, Imitating, and Participating.

Hart, C. W. M. and A. R. Pilling

1960 *The Tiwi of North Australia*. Case Studies in Cultural Anthropology. Holt, Rinehart and Winston, New York.

Henrich, J.

2004 Demography and Cultural Evolution: Why adaptive cultural processes produced maladaptive losses in Tasmania. *American Antiquity* 69(2):197-214.

Henrich, J. and F. 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(3):165-196.

Henrich, N. S. and J. Henrich

in press *Why Humans Cooperate: A Cultural and Evolutionary Explanation* Oxford University Press, Oxford.

Jones, R.

1976 Tasmania: aquatic machines and off-shore islands. In *Problems in Economic and Social Archaeology*, edited by G. Sieveking, I. H. Longworth and K. E. Wilson, pp. 235-263. Duckworth, London.

1977a The Tasmanian paradox. In *Stone tools as cultural markers: change, evolution and complexity*, edited by V. S. Wright, pp. 189-204. Humanities Press, New Jersey.

1977b Why did the Tasmanians stop eating fish? In *Explorations in ethno-archaeology*, edited by R. Gould, pp. 11-47. University of New Mexico Press Alberqueque School of American Research, Santa Fe.

1995 Tasmanian Archaeology: Establishing the Sequence. *Annual Review of Anthropology* 24:423-46.

Kelly, R. L.

1995 *The Foraging Spectrum: Diversity in Hunter-Gatherer Lifeways*. Smithsonian Press, Washington D. C.

Lancy, D. F.

1996 *Playing on Mother Ground: Cultural Routines for Children's Development*. Culture and Human Development. The Guilford Press, London.

McBrearty, S. and A. Brooks

2000 The revolution that wasn't: a new interpretation of the origin of modern human behavior. *Journal of Human Evolution* 39(453-563).

McElreath, R. and R. Boyd

Forthcoming *Modeling the Evolution of Social Behavior*. Princeton University Press, Princeton

Oswalt, W. H.

1976 *An Anthropological Analysis of Food-Getting Technology*. John Wiley, New York.

Radcliffe-Brown, A. R.

1964 *The Andaman Islanders*. The Free Press of Glencoe, New York.

Read, D.

2005 Tasmanian Knowledge and Skill: Maladaptive Imitation or Adequate Technology. *American Antiquity*.

Shennan, S.

2001 Demography and Cultural Innovation: A Model and Its Implications for the Emergence of Modern Human Culture. *Cambridge Archaeology Journal* 11(1):5-16.

Tomasello, M.

1999 *The cultural origins of human cognition*. Harvard University Press, Cambridge, MA.

Ugan, A., J. Bright and A. Rogers

2003 When is technology worth the trouble? *Journal of Archaeological Science* 30:1315-1329.

NOTES

1. I find some ambiguity in Read's presentation of what he thinks I mean. In some places, he says that I have neglected the transmission of knowledge (as an aspect of skill), while in others he says I confounded knowledge and motor skills. That real answer is that both are assumed to be part of the cultural transmission process. But, "confounded" is not accurate, since as I show below, partitioning knowledge and skill does not yield a qualitatively different answer.

2. Mental representations include "motor skills" since they are just as much information stored in individual's nervous system as "knowledge."

3. Readers should also be advised that Read's description of the specifics of my model, including the meaning of the parameters (e.g. N, α), has numerous inaccuracies. It's not possible to understand much of what I did from Read's paper.

4. To see this more sharply, consider that Hamilton's Rule for the evolution of kin-based altruism, $rb > c$ (where r is the coefficient of relatedness, b is the benefit delivered by the altruist and c is the cost of this), requires several unrealistic assumptions. First, to incorporate r into the derivation, one must ignore the effect of natural selection on gene frequencies (odd, since that is precisely what the model is analyzing). Second, the model assumes that genes have additive effects on phenotype (no dominant or recessive genes). Third, the relationship between the benefits delivered by altruists and the fitness gained is assumed to be linear. Usually, in the real world, there is a difference between the benefit you get from the first fish I give you and the 20,000th one—but, not in Hamilton's Rule. All three of these are unrealistic assumptions. But, fear not, they are all also excellent tactical assumptions that do little violence to the selective dynamics governing the evolution of kin-based altruism. Despite these assumptions, Hamilton's Rule has proved incredibly useful to field workers and explains many otherwise puzzling aspects

of nature (McElreath and Boyd forthcoming). Do we reject Hamilton's Rule because of unrealistic assumptions?

5. The degree of interconnectedness among subpopulations will vary for different aspect of culture: e.g., crafts, tools, and medicinal plant knowledge.

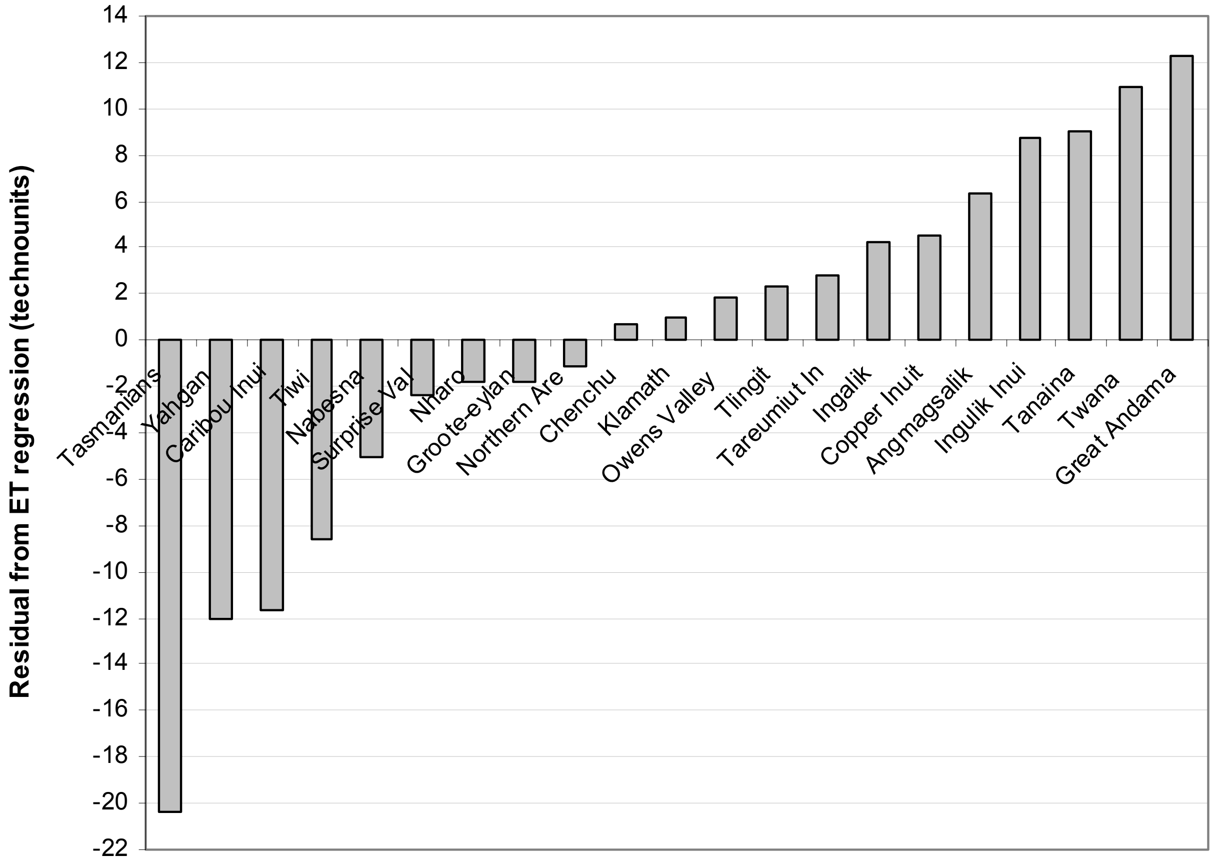
6. Note that this is a stationary analysis, so that the λ values are only locally constant.

7. Similarly for the Paiute, Binford gives us data from 13 Paiute groups, each with quite different population sizes and densities. However, for many aspects of culture (including technology), we probably would not want to assume that Paiute groups, for example, do not transmit to other groups, and all cultural evolution occurs locally—although Read does.

8. As a second check of the density data's robusticity, I compared Read's densities to those found for the same groups in Kelly (1995), matching by ethnolinguistic label. To assess the noise in the data, I calculated the standard deviation in estimated densities for each group, and then calculated the average standard deviation between these data sources. These measures are 18.5 (Kelly), 12.2 (Read), and 10.4 (Read vs. Kelly). The last measure is concerning, since it means that the average standard deviation between Read's and Kelly's estimates is almost as large as the standard deviation in the data themselves: there is an awful lot of noise here—by "noise" I include any temporal and spatial variation that is unaccounted for by the matching ethnolinguistic labels. If one removes the Andamese (a clear outlier in Kelly's data), these standard deviations drop to 8.37, 11.52, and 6.9, respectively. The good news is that the density datasets do correlate 0.58.

9. For statistical reasons, I used the natural logarithm of all of these predictor variables in the regression except number of residential moves per year and percentage contribution of aquatic animals to diet

Figure 1. Bar graph plots the residuals from regressing MXT on ET. The vertical axis gives the difference between each group's actual MXT value and their predicted MXT value.


Residual from ET regression (technounits)
14
12
10
8
6
4
2
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
-20
-22
Tasmanians
Yahgan
Caribou Inui
Tiwi
Nabesna
Surprise Val
Nharo
Groote-eylan
Northern Are
Chenchu
Klamath
Owens Valley
Tlingit
Tareumiut In
Ingalik
Copper Inuit
Angmagsalik
Ingulik Inui
Tanaina
Twana
Great Andama