The Learning and Transmission of Hierarchical Cultural Recipes

Alex Mesoudi
Department of Social and Developmental Psychology
University of Cambridge, UK
am786@cam.ac.uk

Michael J. O'Brien
Department of Anthropology
University of Missouri
Columbia, MO, USA

Abstract

Archaeologists have proposed that behavioral knowledge of a tool can be conceptualized as a "recipe"—a unit of cultural transmission that combines the preparation of raw materials, construction, and use of the tool, and contingency plans for repair and maintenance. This parallels theories in cognitive psychology that behavioral knowledge is hierarchically structured—sequences of actions are divided into higher level, partially independent subunits. Here we use an agent-based simulation model to explore the costs and benefits of *hierarchical* learning relative to *holistic* learning, where entire behavioral sequences are learned in an all-or-nothing fashion, and *diffusionist* learning, where actions are completely independent. Hierarchical learning is favored under the reasonable assumptions that learning is associated with some degree of both error and cost, and that behavior can be grouped into subunits that repeat in one or more tool recipes. These general predictions can be tested in the archaeological and ethnographic record. Recent advances in evolutionary developmental biology have revealed a number of parallels between the hierarchically structured, recipe-like organization of behavioral knowledge that we examine here and the manner in which biological organisms develop.

Keywords

agent-based simulation, behavioral learning, cultural transmission, hierarchical organization, modularity, recipes, scripts, tools

October 4, 2007; revised and accepted May 30, 2008

Cultural Traits as "Recipes"

Two issues of long-standing interest to cultural anthropologists (Lyman and O'Brien 2003), emphasized also by theories of Darwinian cultural evolution (e.g., Cavalli-Sforza and Feldman 1981; Boyd and Richerson 1985, 2005; Mesoudi et al. 2004, 2006), concern the nature of the "cultural trait" and how such traits are learned and culturally transmitted. Whether, and how, culture can be divided into discrete particles or elements has been the subject of much debate, from early 20th century discussion surrounding diffusionism (e.g., Graebner 1911; Benedict 1934) to recent arguments over memes (Bloch 2000; Kuper 2000). The issue of how to delineate cultural traits is also important for applications of phylogenetic methods to cultural data sets (Boyd et al. 1997; O'Brien and Lyman 2003).

For archaeologists, the challenge is how to apply the cultural trait concept to material artifacts such as stone tools, pottery, and projectile technology. Because archaeologists are concerned with the products of past behavior, that is, the artifacts that people constructed and used in the past, their cultural trait concepts tend to focus on behavioral knowledge rather than purely abstract or semantic units of information. The successful construction and use of tools typically involves the execution of a lengthy sequence of actions (Feinman et al. 1981; Schiffer and Skibo 1987; Bleed 2001), from the acquisition and preparation of materials to a tool's eventual use, with each action functionally dependent on previous actions. For example, a tool cannot be used until it has been successfully constructed, and it cannot be constructed without the correct materials, prepared in the correct manner. Consequently, several archaeologists have proposed that a cultural trait can be usefully conceptualized as a "recipe" (Krause 1985; Schiffer and Skibo 1987; Neff 1992; Lyman and O'Brien 2003)—a unit of cultural transmission that combines raw materials and the various behaviors that constitute a person's knowledge regarding how a tool is made and used. A recipe may comprise (1) knowledge of the materials used to construct the tool; (2) instructions for acquiring and preparing those materials; (3) the behavioral knowledge needed to construct the tool; (4) the behavioral knowledge employed in the use of the tool; and (5) contingency behavior for repairing and maintaining the tool.

We make two points in this paper: first, that the recipe concept can be enhanced by recognizing that recipes typically have a hierarchical organization, a claim that is consistent with a large body of work in cognitive psychology and cognitive anthropology; and second, that the reason that behavioral knowledge is often structured in this way is because hierarchical organization facilitates the learning and transmission of complex skills, specifically when learning is associated with some degree of both error and cost, and when behavior can be grouped into subunits that repeat in one or more tool recipes.

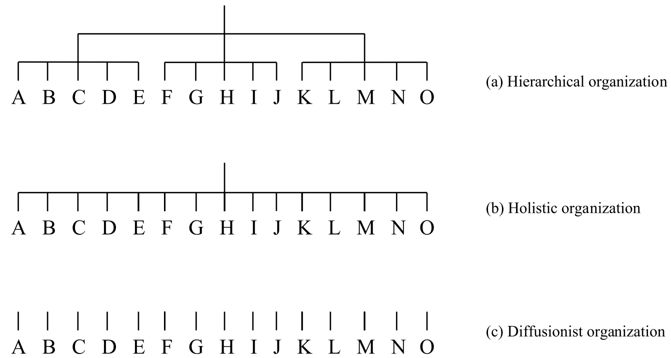

Figure 1.
Three different forms of behavioral organization.

A simple agent-based model is used to formalize and test the latter point.

Recipes Are Hierarchical

We suggest that the recipe concept described above can usefully be seen as *hierarchically structured* (Figure 1a), with the tool trait comprising several behavioral subroutines (e.g., preparation of material, production, use), each of which in turn can be subdivided into a sequence of constituent lower level actions required to complete each subroutine. To clarify this concept, we also define two other forms of organization with which hierarchy can be contrasted. *Holistic* organization (Figure 1b) is where the entire cultural tool trait is viewed as a single entity or continuum and, while division into separate actions is possible, these actions are not grouped into subroutines, nor can they be performed or learned separately as part of other tool traits. *Diffusionist* organization (Figure 1c) is where the tool trait is seen to be composed of a series of independent and ungrouped actions each of which can be learned and performed separately, perhaps as part of other traits. These distinctions will become clearer when we present the agent-based model.

Considerable evidence from cognitive psychology, cognitive anthropology, and other fields supports the notion that skills, such as those involved in tool making, are represented, learned, and transmitted hierarchically, rather than in a holistic or diffusionist manner. Cognitive psychologists Schank and Abelson (1977) proposed that behavioral knowledge related to routine events is structured hierarchically (Figure 1a). For example, "going to a restaurant" entails the completion of a number of subroutines, such as "sit down," "order food," "eat food," "pay bill," and "leave." Each in turn comprises a series of lower level actions, such as "pick up menu," "read menu," "decide what to eat," and so on. Memory studies (e.g., Bower et al. 1979; Abbott et al. 1985; Zacks and Tversky 2001; Zacks et al. 2001) have consistently found that people (1) agree on how to group lower level actions into higher level subunits; (2) spontaneously reintroduce the hierarchical structure of scrambled scripts; and (3) readily infer absent lower level actions from higher level subunits, all consistent with a common underlying hierarchical organization of behavioral knowledge. Studies from developmental psychology (e.g., Slackman et al. 1986; Bauer and Mandler 1989; van den Broek et al. 1996) show that 2- to 6-year-old children develop an increasingly elaborate hierarchical organization in their understanding of behavioral routines, with both the number of elements (e.g., subunits or actions) in the hierarchy, and the children's understanding of how those elements fit into the hierarchy increasing with age and experience.

There is also evidence that cultural transmission, like individual learning, is guided by hierarchical constraints. Whiten (2002) found that 3 year olds copied the hierarchical organization of a behavioral routine (opening an artificial fruit) but not the lower level sequences of actions, and Mesoudi and Whiten (2004) found evidence for the spontaneous introduction of hierarchical organization as written descriptions of script-like behavior were passed along chains of adult participants. Byrne and Russon (1998), meanwhile, present observational evidence that nonhuman primates imitate behavior hierarchically. For example, mountain gorillas tend to imitate behaviors at the subgoal level, rather than at the level of overall goals or fine motor actions.

Cognitive anthropologists and archaeologists have obtained similar evidence for the hierarchical organization of behavioral knowledge in the ethnographic and archaeological record. Gatewood (1985), in an ethnographic study of Alaskan salmon fishermen, found that the behavioral knowledge of novice fishermen initially comprised a sequence of unconnected and unorganized actions (akin to a "string of beads" in the words of Gatewood, and resembling our "diffusionist" organization: Figure 1c). As fishermen gained experience, however, this knowledge became increasingly hierarchically organized into distinct subroutines, for example, "making a set, pursing, hauling gear" (Gatewood 1985: 207). Keller and Keller (1996) similarly found that the tool-making knowledge of expert blacksmiths was hierarchically organized, such that separate actions were grouped into a series of segments (or "constellations"), all within an overall "umbrella plan." Stout (2002) reports a similar hierarchical segmentation in the behavioral knowledge of stone-tool knappers in Indonesia, with actions separated into higher level subroutines such as "raw-material procurement, roughing-out, grinding, and so on" (Stout 2002: 705). Interestingly, whereas Gatewood (1985) found that the knowledge of novice fishermen comprised sequences of unconnected and unorganized actions, Stout (2002: 708) found that novice knappers showed a more global or holistic behavioral strategy with no higher level organization at all (resembling our "holistic" organization: Figure 1b).

Finally, in an analysis of Late Paleolithic Japanese stone tools, Bleed (2002) found evidence for the hierarchical organization of stone tool making, with sequences of actions separated into higher order segments, such that interruptions in the behavioral sequences most often occurred at the end of one segment and before the next began. In his conclusion, Bleed (2002: 342) echoes our own reading of the literature in stating that stoneworkers "were guided by conceptualizations that made them conceive of their task as something other than either a simple continuum (i.e., holistic organization) *or* a sequence of discreet steps (i.e., diffusionist organization)." Ethnographic and archaeological evidence, then, generally suggests that the behavioral knowledge of novices tends to be either (1) a sequence of unconnected, separate steps, as in Gatewood's (1985) novice fishermen, with no higher level organization, or (2) guided by a global strategy with no lower level organization, as in Stout's (2002) novice knappers. Experienced individuals, by contrast, combine (1) and (2) to form (3), hierarchically organized behavioral knowledge, with actions grouped into subroutines, and subroutines into global strategies.

Why Hierarchical Organization?

The ethnographic evidence discussed in the previous section suggests that the behavioral knowledge of novices becomes increasingly hierarchically organized as the novice masters the skill, either by practice (individual learning) or by imitation and teaching (cultural transmission). This suggests that hierarchical organization has some benefits over the other forms of organization—holistic and diffusionist—that were observed in novices. What might the functional advantage(s) of hierarchical organization be? Simon (1962) argued that hierarchical organization should be favored over holistic organization because the former provides a series of stable subassemblies, such that any error or interruption to a hierarchical sequence affects only the current subunit without disrupting subunits that have already been learned or executed. Interruptions or errors in sequences that are holistically organized, on the other hand, would disrupt the entire sequence of all actions. Hence, hierarchically structured behavioral sequences should be less vulnerable to error and more easily executed and learned than holistically structured behavioral sequences.

Although this advantage would favor hierarchical organization over holistic organization, diffusionist organization, where each behavioral element in a sequence is acquired or executed independently, would appear to be favored over both of these. Any error in the learning of a diffusionist sequence would affect only the current action and not disrupt the acquisition/execution of any other behavior. In effect, diffusionist organization can be seen as an extreme form of hierarchical organization, where each subunit comprises a single action. Are there any reasons we might expect hierarchical learning to be favored over diffusionist learning? Lyman and O'Brien

(2003) suggest that if cultural traits have a recipe-like hierarchical organization, then the different subunits can be rearranged to form different recipes (similar to Simon's [1962] "redundancy" argument). Hence, a single subunit can be repeatedly used and benefited from, yet has to be learned and mastered only once. Repetition of subunits would reduce the overall cost, in terms of time and effort, of learning and executing multiple tool recipes. A diffusionist learner would not benefit from this repetition, as the elements of behavior are not represented as subunits that can be repeated.

We now introduce a simple agent-based simulation model to formally specify and test these two arguments regarding the relative advantages and disadvantages of holistic, hierarchical, and diffusionist learning. Note that this model is intended primarily as a means of formalizing and testing our arguments outlined above, and not a direct simulation of human behavior.

The Agent-Based Model

The agent-based model features successive generations of computer-generated agents, each of which individually learns a series of recipe-like sequences of actions (cultural transmission is dealt with in the following section; here we focus on individual learning). Agents can acquire knowledge of five separate *recipes,* each of which consists of a sequence of 25 *actions*. These actions, depending on the agent's learning rule (see below), can be subdivided into five *subunits,* each comprising five actions. An agent's knowledge of an action is represented by a single value ranging between 0 and 1 to an accuracy of three decimal places. Each action also has a randomly generated optimal value between 0 and 1. During individual learning, each agent cycles through each action in order and successfully learns the optimal value with probability ($1 - L_{error}$), where L_{error} represents error, or inaccuracy in learning. Hence, when $L_{error} = 0$, learning is perfect and error free, whereas values greater than 0 represent increasingly inaccurate and error-prone learning.

Each agent has L_t trials of learning, with one learning attempt (whether successful or unsuccessful) per trial. Each trial incurs a cost of L_{cost}, whether successful or unsuccessful. Failure to learn a behavior has different consequences depending on the agent's learning rule. Holistic (*Hol*) agents must successfully learn the entire behavioral sequence of a recipe (25 successful learning trials); otherwise, they lose knowledge of the recipe and must start again from the first action. Hierarchical (*Hier*) agents must successfully learn the behavioral sequence of the current subunit only. Failure to learn a behavior means that *Hier* agents lose knowledge of that subunit and must start the subunit again, with already completed subunits left unaffected. Finally, diffusionist (*Diff*) agents learn each action independently, and failure to learn an action does not affect any previously learned action. Hence, we can test the arguments that *Hier* agents outperform *Hol* agents as a result of the stability of subunits in hierarchical learning, and that *Diff* agents outperform both *Hol* and *Hier* agents because errors in diffusionist learning do not negatively affect the learning of any other action. These effects should occur only when $L_{error} > 0$.

To test these predictions, agents are replaced after each generation with a new generation of agents, who inherit their learning rule from their parents. This allows more successful rules to increase in frequency. There is a small probability ($m = 0.01$) that the learning rule mutates into a random alternative. In the first generation, the learning rule (*Hol, Hier* or *Diff*) is assigned at random. Agents also inherit L_t but not the values of the actions. For simplicity we assume that reproduction is asexual and generations are discrete. Agents reproduce with a probability equal to their fitness, which is the proportion of the 125 actions that have been successfully learned. We assume that each recipe is functionally interlinked, such that if any one of the 25 actions in a recipe does not match the optimal value, then the entire recipe yields a fitness of zero. The model was run with a population of 200 agents over 10,000 generations. Results given are the averages of 10 independent runs of 10,000 generations. The model was analyzed by systematically varying each parameter across a range of representative values at each level of every other parameter.

Simulation Results

First we consider the case where only *Hier* and *Hol* agents are present in the population. Figure 2a shows that when $L_{error} = 0$, neither *Hier* nor *Hol* agents are favored and that when $L_{error} > 0$, *Hier* agents are favored, as predicted. However, this effect is moderated by the number of learning trials, L_t. When L_t is large, *Hier* agents no longer have an advantage over *Hol* agents because there are enough learning trials for *Hol* agents to successfully learn the recipes despite the error in learning. Because it seems unrealistic that the length of learning, L_t, would increase indefinitely, we can set $L_{cost} > 0$ and allow L_t to mutate with a small probability (with a probability of $p = 0.01$, the L_t of offspring randomly either increase or decrease by one unit from their parents' L_t). When $L_{cost} > 0$ and $L_{error} > 0$, *Hier* learning is now selected and maintained at equilibrium for all starting values of L_t because the cost of learning drives L_t down to a minimum value, at which *Hier* agents, but not *Hol* agents, can learn all the recipes. The time series in Figure 2b shows a typical run where $L_{cost} > 0$ and $L_{error} > 0$.

We now consider the case where only *Hier* and *Diff* agents are present. (Simulations were also run with all three types of agents, but in no case did *Hol* agents outperform either *Hier* or *Diff*. For the convenience of presentation, results are given for populations with only *Hier* and *Diff* agents.) Error in learning

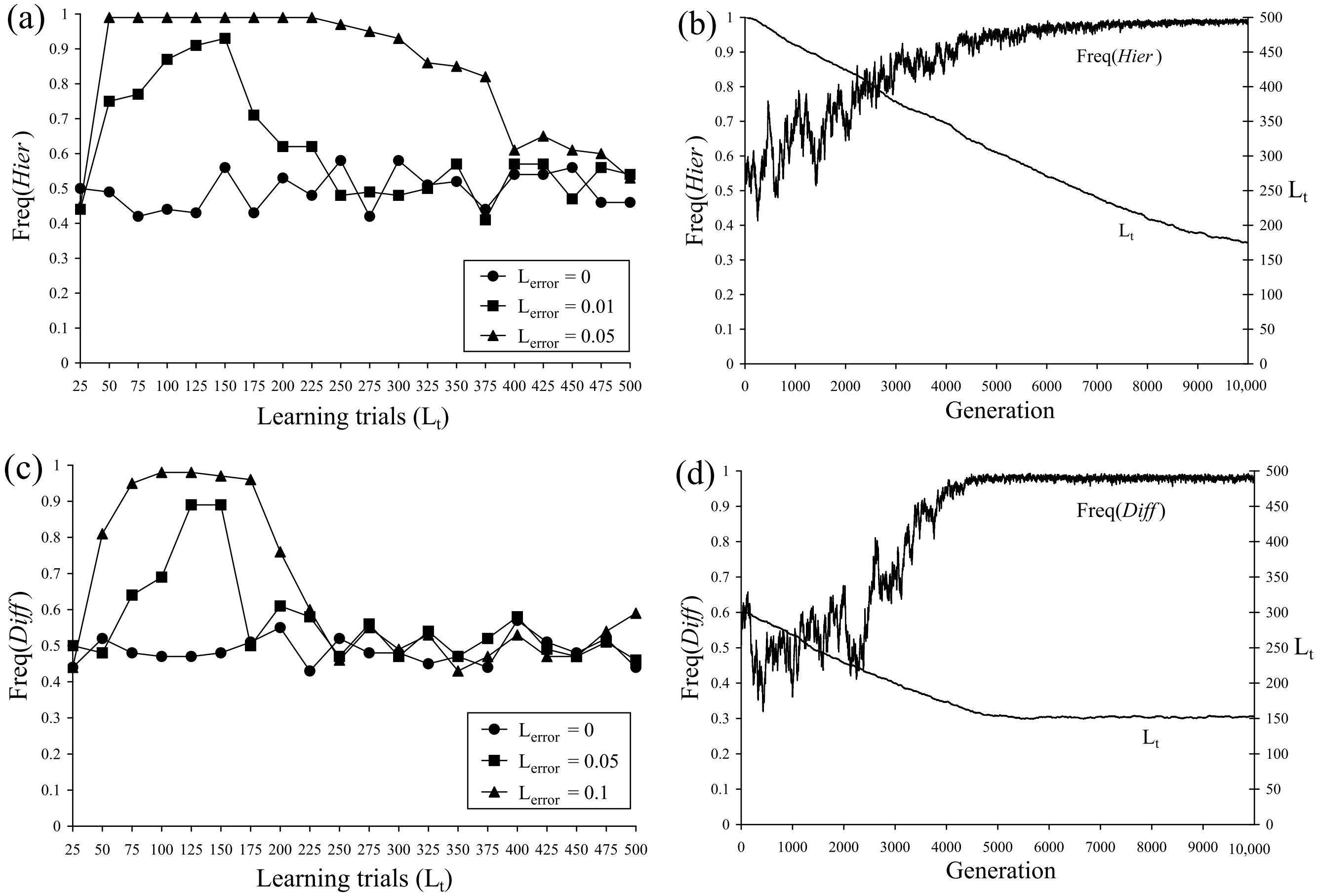

Figure 2.
(a) When learning is associated with error ($L_{\text{error}} > 0$) but no cost ($L_{\text{cost}} = 0$), *Hier* agents are favored over *Hol* agents, but only when there are relatively few learning trials (L_t); (b) when learning is both costly ($L_{\text{cost}} = 0.001$) and imperfect ($L_{\text{error}} = 0.05$), there is a reduction in L_t, which maintains the advantage of *Hier* agents over *Hol* agents at all values of L_t; (c) and (d) show similar patterns when *Diff* agents compete with *Hier* agents.

($L_{\text{error}} > 0$) favors *Diff* agents over *Hier* agents, as predicted, and this effect is again moderated by L_t (Figure 2c). At large values of L_t, *Hier* agents can learn all recipes despite the error in learning. Introducing a cost to learning ($L_{\text{cost}} > 0$) again selects against L_t and maintains the advantage to *Diff* agents at any starting value of L_t (e.g., Figure 2d). This effect (as well as the advantage of *Hier* over *Hol* agents noted above) is observed for all values of L_{error} greater than zero and below an upper limit—that is, when learning is neither perfect nor impossible. Above the upper limit, learning is too inaccurate for agents to learn any of the recipes, and the lack of differential fitness prevents the selection of learning rules.

To test the second argument—that hierarchical learning is favored when subunits repeat in more than one recipe—we introduce a probability, r, that the optimal values of all actions in a single subunit are copied from the optimal values of a random subunit from the previous recipe. Hence, when $r = 1$, all recipes are identical; when $r = 0.5$, each recipe shares on average half of its subunits with the previous recipe; and when $r = 0$, no recipe shares any subunit (as already considered above). *Hier* agents can uniquely take advantage of this subunit repetition during learning. If the first action of the current subunit matches the first action of any of the subunits that a *Hier* agent has previously successfully learned, the agent automatically learns the entire subunit in a single round of learning. Hence, hierarchical learners can learn all five recipes in fewer rounds of learning than the other two types of learners, and therefore bear less fitness cost (when $L_{\text{cost}} > 0$) and suffer less from the disruptive effects of learning error (when $L_{\text{error}} > 0$). Simulations confirmed that *Hier* agents are indeed favored when r is sufficiently large and only when $L_{\text{cost}} > 0$ and/or $L_{\text{error}} > 0$ (Figure 3a).

The probability, r, also interacts with the number of subunits, the number of actions per subunit, and the number of recipes (Figure 3b). Increasing the number of recipes increases the overall number of repeated subunits, thereby increasing the benefit to *Hier* agents. Increasing the number of subunits per recipe has the same effect as increasing the number of repeated subunits and thereby favors *Hier*, until there are too many subunits to learn an entire recipe given the error in learning. The number of actions per subunit favored *Hier* agents at relatively low values, as subunits that contained too many actions were

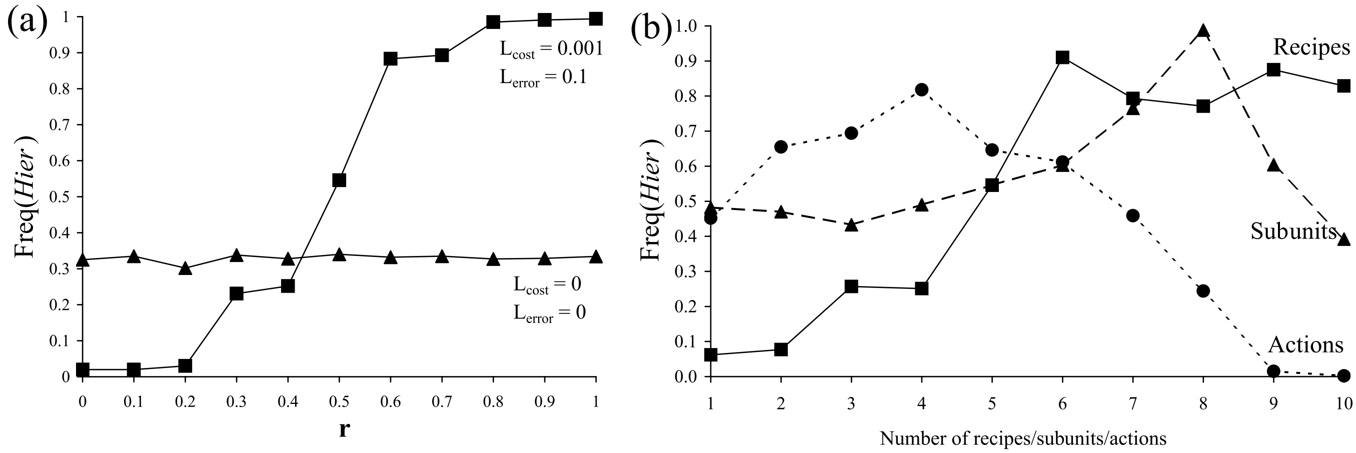

Figure 3.
(a) Repeating subunits ($r > 0$) favor *Hier* learners when learning is costly and imperfect. Frequencies are the means of 10 runs after 10,000 generations. All three types of agents initially present, although the final frequency of *Hol* was never greater than 0.01. (b) The effect of varying the number of recipes, subunits per recipe, and actions per subunit on the frequency of *Hier* agents. All other agents were *Diff*; parameters $L_{\text{cost}} = 0.001$, $L_{\text{error}} = 0.1$, $r = 0.5$.

too difficult to learn, thereby negating the original advantage of hierarchical organization.

A further factor to be considered is change in the selective environment. This change can be *action specific*, where the optimal value associated with a single randomly chosen action is changed to a new random number between 0 and 1; *subunit specific*, where the optimal values of every action within a single randomly chosen subunit are changed; or *recipe specific*, where the optimal values of an entire randomly chosen recipe are changed. In the latter two forms of change, r still determines the probability that a changed subunit is repeated from a previous recipe. Action-specific, subunit-specific, and recipe-specific change occurs in each generation with a probability of c_1, c_2, and c_3, respectively.

Figure 4a shows that a small probability of action-specific environmental change ($c_1 = 0.01$, $c_2 = 0$, $c_3 = 0$) causes diffusionist learning to supplant hierarchical learning. This is because *Hier* learners must relearn the entire subunit in which change occurs, whereas *Diff* learners suffer no wider negative effects beyond the single behavior that has changed. Moreover, any change within a subunit that has been repeated from the previous recipe (given $r > 0$) will mean that the repeated subunits are no longer identical, thereby preventing *Hier* agents from benefiting from subunit repetition. However, this kind of isolated environmental change seems unlikely, given that we have assumed that the actions are functionally linked. With respect to actual recipes, a change in the available raw materials is likely to necessitate a change in the preparation of that material. Similarly, a change in how a tool is used would likely necessitate some change in its construction to accommodate the change in use. We therefore also consider the environmental change that is subunit specific ($c_1 = 0$, $c_2 = 0.01$, $c_3 = 0$) and recipe specific ($c_1 = 0$, $c_2 = 0$, $c_3 = 0.01$). Both of these maintain *Hier* agents in the population (Figure 4(a)).

Extension of the Model: Vertical Cultural Transmission

Contrary to the preceding model, technologies and practices are not always reinvented anew each generation; rather, they typically are learned from other members of society. Ethnographic studies of modern nonindustrial peoples suggest that functionally interlinked, recipe-like behavioral knowledge is acquired from others through a lengthy period of instruction and (primarily) observation (Schiffer and Skibo 1987; VanPool et al. 2008). Given such a lengthy period of learning, recipe-like behavior is most likely to be acquired from parents, with whom offspring not only spend most of the time and have more opportunity to observe, but also have a genetic interest in successfully and accurately passing on subsistence and technological knowledge (thereby enhancing inclusive fitness; Hamilton 1964). This is consistent with anthropological evidence that cultural transmission is predominantly vertical in many traditional societies for many traits (Hewlett and Cavalli-Sforza 1986; Guglielmino et al. 1995; Ohmagari and Berkes 1997; Hewlett et al. 2002; O'Brien et al. 2008), including specific ethnoarcheological evidence for the vertical transmission of material culture (Neff 1992; VanPool et al. 2008). We therefore predict that functionally linked behavioral sequences will be vertically transmitted when vertical transmission is less costly and/or features less error than individual learning, both of which seem realistic assumptions. The general findings from the previous section regarding individual learning should also apply to vertical transmission—that is, error and cost in vertical transmission should favor diffusionist vertical

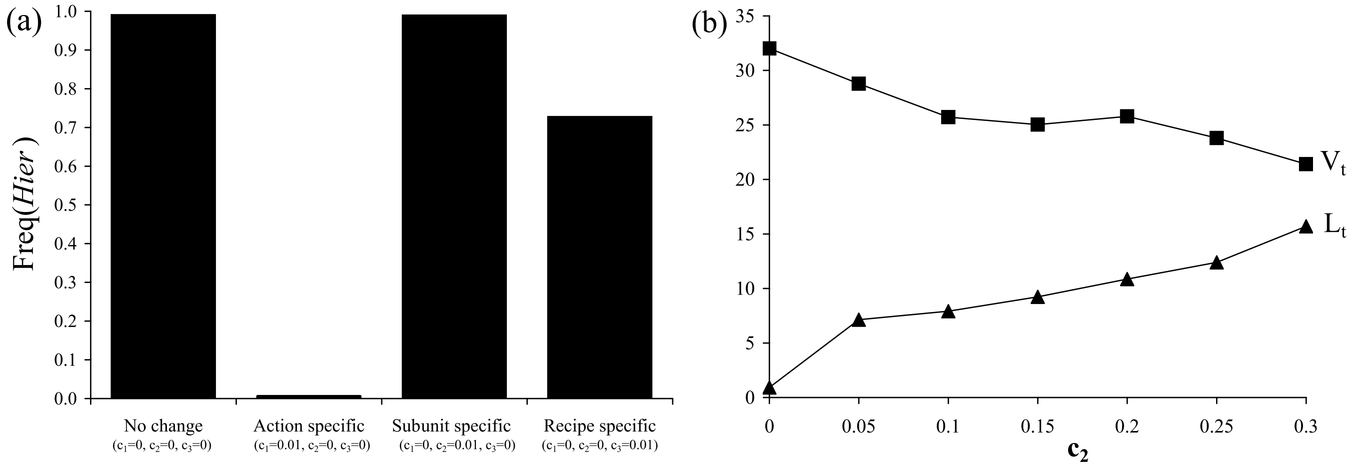

Figure 4.
(a) Effect of type of environmental change on the frequency of *Hier* agents (frequencies are means of 10 runs after 10,000 generations; all other agents *Diff*; $L_{error} = 0.1$, $L_{cost} = 0.001$, $r = 0.75$); (b) increasing the probability of environmental change—in this case, subunit specific ($c_2 > 0$)—increases the length of individual learning (L_t) and decreases the length of vertical transmission (V_t).

transmission whereas repetition of subunits should favor hierarchical vertical transmission.

We introduce in the agent-based model the possibility of vertical transmission of behavior from parent to offspring in order to explore the conditions under which individual learning is replaced with vertical cultural transmission and the form that this vertical transmission takes. We are interested in changes in the number of trials of vertical transmission (V_t) and the number of trials of individual learning (L_t), and in the learning rules used during these periods (*Hier, Hol*, or *Diff*). Vertical transmission occurs identically to individual learning except that the parent's values are copied rather than the optimal values. The parent's values are successfully copied with a probability ($1 - V_{error}$), where V_{error} represents error in vertical transmission. There are V_t rounds of vertical transmission, one per copying attempt (whether successful or unsuccessful), each imposing a fitness cost of V_{cost}. Vertical transmission may be *Hol*, *Hier*, or *Diff*, each of which has the same consequences as described above for individual learning. Agents may have different learning rules for individual learning and for vertical transmission. There is no vertical transmission in the first generation, as there is no parental generation.

Simulations showed that the form of learning (vertical cultural transmission or individual learning) that is less costly dominates when error is kept constant (Figure 5a). Similarly, the form of learning that is more accurate dominates when cost is kept constant (Figure 5b). When vertical transmission is the predominant form of learning, the parameters V_{error}, V_{cost}, r, c_1, c_2, and c_3 have identical effects on the frequencies of agents who have *Hol, Hier,* or *Diff* vertical transmission rules as L_{error}, L_{cost}, r, c_1, c_2, and c_3 do on the relative frequencies of *Hol, Hier*, and *Diff* individual learning rules. Specifically, the same three key findings were observed for vertical transmission as for individual learning: (1) when $V_{error} > 0$, *Hier* transmission is favored over *Hol* transmission and *Diff* transmission is favored over *Hier*/*Hol* transmission, as long as $V_{cost} > 0$, thereby keeping V_t to a minimum; (2) when $r > 0$, *Hier* transmission is favored over *Diff*/*Hol* transmission; and (3) when environmental change is action specific ($c_1 > 0$), *Diff* transmission is favored over *Hier*/*Hol* transmission, and when environmental change is either subunit specific ($c_2 > 0$) or recipe specific ($c_3 > 0$), *Hier* transmission is favored over *Diff*/*Hol* transmission. Environmental change also increases the length of individual learning (L_t) when vertical transmission is predominant (Figure 4b) because vertical transmission alone cannot track novel environmental change.

Discussion

Tools can be seen as the products of lengthy sequences of functionally linked actions, from the preparation of raw materials to the skills involved in a tool's eventual use. We have drawn on concepts and arguments from cognitive psychology and the behavioral sciences to explore the relative costs and benefits of three different ways of organizing the behavioral knowledge involved in the learning and transmission of such behavioral sequences. A simple agent-based simulation model was used to support the arguments that hierarchically organized learning is favored over holistic, all-or-nothing learning when there is some degree of error in learning. This is because subunits provide stable intermediate stages. This advantage is maintained at equilibrium when learning is associated with some cost, which keeps the amount of time spent learning to a minimum. That learning exhibits both cost and error seems a realistic assumption, given that mastering the skills required to make

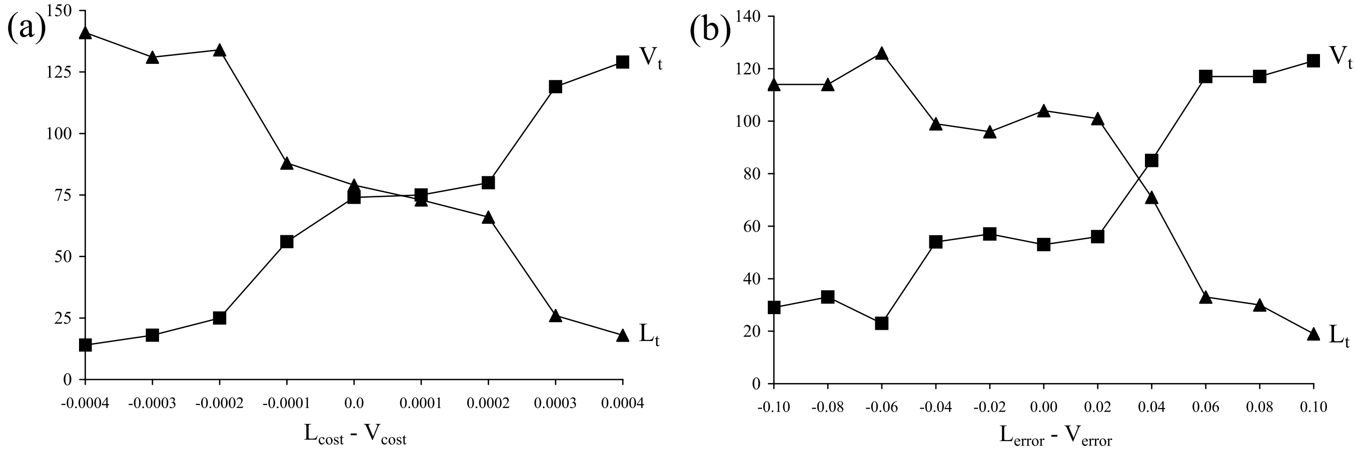

Figure 5.
(a) The number of trials of the less costly form of learning (vertical transmission or individual learning) increases in frequency, with error kept constant ($L_{\mathrm{error}} = V_{\mathrm{error}} = 0.1$), as does (b) the more accurate form of learning, with cost kept constant ($L_{\mathrm{cost}} = V_{\mathrm{cost}} = 0.001$).

and use tools typically requires repeated practice over several years (Kramer 1985; Stout 2002). Hierarchical learning is favored over diffusionist learning—in which actions are learned independently in a piecemeal fashion—when subunits repeat in one or more recipe. This is because the overall cost of learning is reduced. Hierarchical learning is more likely to emerge when there are many repeating subunits (e.g., when there are multiple recipes with multiple subunits and few actions per subunit) and when environmental change affects entire subunits and/or entire recipes (consistent with a functionally interlinked recipe).

Finally, the vertical cultural transmission of behavioral knowledge from the previous generation is more likely to replace individual learning when the former is less costly and features less error. This assumption is consistent with both theoretical predictions (the maximization of inclusive fitness) and ethnographic evidence. Some degree of individual learning is retained when the selective environment changes, which vertical transmission alone cannot track. A plausible scenario suggested by our model is, therefore, one in which there is an extended period of relatively low cost and relatively accurate vertical cultural transmission where hierarchically structured behavioral knowledge is learned from the parental generation, along with less frequent individual learning that is predominantly diffusionist, that is, single actions are learned independently of other actions, and functions to track novel environmental change.

Several of these assumptions and predictions can be tested with further archaeological and ethnographic study. If lengthy sequences of functionally interlinked behaviors such as those behind tool construction and use are transmitted predominantly vertically, then phylogenetic analyses may be useful for detecting such traditions (O'Brien and Lyman 2000, 2003). Ethnographic studies of technology learning might provide formal, quantitative analyses of the relative cost and accuracy of individual learning and cultural transmission, given that these factors strongly determine the nature of learning in our model (hierarchical or diffusionist; individual or cultural). Hierarchical organization was found to be advantageous only when subunits repeat in one or more recipes, as these repeated subunits have to be learned only once and so reduce the overall costs of learning. Given our earlier discussion of the prevalence of hierarchical learning, we might expect to find evidence of repeated subunits in the archaeological record, for example, where the same technological component is repeated in a single tool.

Not considered here is division of labor, which ethnographic evidence suggests is common in groups that manufacture tools (Kramer 1985). Division of labor might further promote hierarchical learning, with different members of a group learning and specializing in different subunits, thereby achieving faster and more accurate learning than if everyone had to learn the entire recipe. This cooperation would bring with it a vulnerability to free riders, however, given that different subunits are unlikely to be equally easy to learn (Feinman et al. 1981).

Recent advances in evolutionary developmental biology or EvoDevo (e.g., Carroll 2005), have revealed a number of parallels between the hierarchically structured, recipe-like organization of behavioral knowledge that we have examined here and the manner in which biological organisms develop (Schlosser and Wagner 2003; Callebaut and Rasskin-Gutman 2005). Phenotypic characters are often modular (Wagner and Altenberg 1996; Hansen 2003) such that different characters develop as partially self-contained modules, similar to the subunits of a behavioral recipe. These modules are organized hierarchically, with a small number of higher level regulatory genes triggering the growth of entire lower level modules, such as the *Drosophila* "eyeless" gene that triggers the development of an entire eye in different parts of the body (Halder et al. 1995) or the *Hox* genes that control the growth of limbs or body segments (Carroll 1995; Meyer 1998). Consequently, bodies can be built by repeating modular body parts, such as limbs, teeth, or body segments (Weiss 1990), in the same way in which behavioral subunits can be repeated in one or more recipes. These parallels suggest that the advantages of hierarchical organization—localization of error and repetition of subunits—are likely to generalize to many or all knowledge-gaining evolutionary systems (Simon 1962; Dawkins 1976).

References

Abbott V, Black JB, Smith EE (1985) The representation of scripts in memory. Journal of Memory and Language 24: 179–199.

Bauer PJ, Mandler JM (1989) One thing follows another: Effects of temporal structure on 1- to 2-year-olds' recall of events. Developmental Psychology 25: 197–206.

Benedict R (1934) Patterns of Culture. Boston: Houghton Mifflin.

Bleed P (2001) Trees or chains, links or branches: Conceptual alternatives for consideration of stone tool production and other sequential activities. Journal of Archaeological Method and Theory 8: 101–127.

Bleed P (2002) Obviously sequential, but continuous or staged? Refits and cognition in three late paleolithic assemblages from Japan. Journal of Anthropological Archaeology 21: 329–343.

Bloch M (2000) A well-disposed social anthropologist's problems with memes. In: Darwinizing Culture (Aunger R, ed), 189–204. Oxford: Oxford University Press.

Bower GH, Black JB, Turner TJ (1979) Scripts in memory for text. Cognitive Psychology 11: 177–220.

Boyd R, Borgerhoff Mulder M, Durham WH, Richerson PJ (1997) Are cultural phylogenies possible? In: Human by Nature: Between Biology and the Social Sciences (Weingart P, Mitchell SD, Richerson PJ, Maasen S, eds), 355–386. Mahwah, NJ: Erlbaum.

Boyd R, Richerson PJ (1985) Culture and the Evolutionary Process. Chicago: University of Chicago Press.

Boyd R, Richerson PJ (2005) The Origin and Evolution of Cultures. Oxford: Oxford University Press.

Byrne RW, Russon AE (1998) Learning by imitation: A hierarchical approach. Behavioral and Brain Sciences 21: 667–721.

Callebaut W, Rasskin-Gutman D, eds (2005) Modularity: Understanding the Development and Evolution of Natural Complex Systems. Cambridge, MA: MIT Press.

Carroll RL (2005) Endless Forms Most Beautiful: The New Science of Evo-Devo and the Making of the Animal Kingdom. New York: Norton.

Carroll SB (1995) Homeotic genes and the evolution of arthropods and chordates. Nature 376: 479–485.

Cavalli-Sforza LL, Feldman MW (1981) Cultural Transmission and Evolution: A Quantitative Approach. Princeton, NJ: Princeton University Press.

Dawkins R (1976) Hierarchical organisation: A candidate principle for ethology. In: Growing Points in Ethology (Bates PP, Hinde RA, eds), 7–54. Cambridge: Cambridge University Press.

Feinman GM, Upham S, Lightfoot KG (1981) The production step measure: An ordinal index of labor input in ceramic manufacture. American Antiquity 46: 871–884.

Gatewood J (1985) Actions speak louder than words. In: Directions in Cognitive Anthropology (Dough J, ed), 199–219. Urbana, IL: University of Illinois Press.

Graebner F (1911) Methode der Ethnologie. Heidelberg, Germany: Ethnologische Bibliothek.

Guglielmino CR, Viganotti C, Hewlett B, Cavalli-Sforza LL (1995) Cultural variation in Africa: Role of mechanisms of transmission and adaptation. Proceedings of the National Academy of Sciences of the USA 92: 585–589.

Halder G, Callaerts P, Gehring WJ (1995) Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila. Science 267: 1788–1792.

Hamilton WD (1964) The genetical evolution of social behaviour I and II. Journal of Theoretical Biology 7: 1–52.

Hansen TF (2003) Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolvability. BioSystems 2189: 1–12.

Hewlett B, Cavalli-Sforza LL (1986) Cultural transmission among Aka pygmies. American Anthropologist 88: 922–934.

Hewlett B, De Silvestri A, Guglielmino CR (2002) Semes and genes in Africa. Current Anthropology 43: 313–321.

Keller CM, Keller JD (1996) Cognition and tool use: The blacksmith at work. Cambridge: Cambridge University Press.

Kramer C (1985) Ceramic ethnoarchaeology. Annual Review of Anthropology 14: 77–102.

Krause RA (1985) The Clay Sleeps. Tuscaloosa, AL: University of Alabama Press.

Kuper A (2000) If memes are the answer, what is the question? In: Darwinizing Culture (Aunger R, ed), 175–188. Oxford: Oxford University Press.

Lyman RL, O'Brien MJ (2003) Cultural traits: Units of analysis in early twentieth-century anthropology. Journal of Anthropological Research 59: 225–250.

Mesoudi A, Whiten A (2004) The hierarchical transformation of event knowledge in human cultural transmission. Journal of Cognition and Culture 4: 1–24.

Mesoudi A, Whiten A, Laland KN (2004) Is human cultural evolution Darwinian? Evidence reviewed from the perspective of *The Origin of Species*. Evolution 58: 1–11.

Mesoudi A, Whiten A, Laland KN (2006) Towards a unified science of cultural evolution. Behavioral and Brain Sciences 29: 329–383.

Meyer A (1998) *Hox* gene variation and evolution. Nature 391: 225–228.

Neff H (1992) Ceramics and evolution. Archaeological Method and Theory 4: 141–194.

O'Brien MJ, Lyman RL (2000) Applying Evolutionary Archaeology: A Systematic Approach. New York: Kluwer Academic/Plenum.

O'Brien MJ, Lyman RL (2003) Cladistics and Archaeology. Salt Lake City, UT: University of Utah Press.

O'Brien MJ, Lyman RL, Collard M, Holden CJ, Gray RD, Shennan SJ (2008) Phylogenetics and the evolution of cultural diversity. In: Cultural Transmission and Archaeology: Issues and Case Studies (O'Brien MJ, ed), 39–58. Washington, DC: Society for American Archaeology Press.

Ohmagari K, Berkes F (1997) Transmission of indigenous knowledge and bush skills among the Western James Bay Cree women of subarctic Canada. Human Ecology 25: 197–222.

Schank RC, Abelson RP (1977) Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures. Oxford: Erlbaum.

Schiffer MB, Skibo JM (1987) Theory and experiment in the study of technological change. Current Anthropology 28: 595–622.

Schlosser G, Wagner GP, eds (2003) Modularity in Development and Evolution. Chicago: University of Chicago Press.

Simon HA (1962) The architecture of complexity. Proceedings of the American Philosophical Society 106: 467–482.

Slackman EA, Hudson JH, Fivush R (1986) Actions, actors, links and goals: The structure of children's event representations. In: Event Knowledge: Structure and Function in Development (Nelson K, ed), 47—69. Hillsdale, NJ: Erlbaum.

Stout D (2002) Skill and cognition in stone tool production. Current Anthropology 43: 693–722.

van den Broek P, Lorch EP, Thurlow R (1996) Children's and adults' memory for television stories: The role of causal factors, story-grammar categories, and hierarchical level. Child Development 67: 3010–3028.

VanPool TL, Palmer CT, VanPool CS (2008) Horned serpents, tradition, and the tapestry of culture. In: Cultural Transmission and Archaeology: Issues and Case Studies (O'Brien M, ed). Washington, DC: Society for American Archaeology.

Wagner GP, Altenberg L (1996) Complex adaptations and the evolution of evolvability. Evolution 50: 967–976.

Weiss KM (1990) Duplication with variation: Metameric logic in evolution from genes to morphology. American Journal of Physical Anthropology 33: 1–23.

Whiten A (2002) Imitation of sequential and hierarchical structure in action: Experimental studies with children and chimpanzees. In: Imitation in Animals and Artifacts: Complex Adaptive Systems (Dautenhahn K, Nehaniv CL, eds), 191–209. Cambridge, MA: MIT Press.

Zacks JM, Tversky B (2001) Event structure in perception and conception. Psychological Bulletin 127: 3–21.

Zacks JM, Tversky B, Iyer G (2001) Perceiving, remembering, and communicating structure in events. Journal of Experimental Psychology: General 130: 29–58.