Social foraging: individual learning and cultural transmission of innovations

We present two stochastic models of individual and social learning that count the number of individuals exhibiting a learned, resource-producing trait in a group of social foragers. The novelty of our modeling results from incorporating the empirically based assumption that rates of both individual and social learning should depend on the frequency of the learned trait within the group. When resources occur as clumps shared by group members, a naive individual's acquisition of the skill required for clump discovery/production should involve opposing processes of frequency dependence. The opportunity to learn via cultural transmission should increase with the trait's frequency, but the opportunity for learning individually should decrease as the trait's frequency increases. The results of the model suggest that the evolution of the capacity for cultural transmission may be promoted in environments where scrounging at resource clumps inhibits rates of individual learning. *Key words:* cultural transmission, social foraging, social learning. *[Behav Ecol 5:35–43 (1994)]*

Luc-Alain Giraldeau
Department of Biology,
Concordia University,
1455 ouest, Boulevard de Maisonneuve,
Montréal, Québec
H3G 1M8, Canada

Thomas Caraco
Department of Biological Sciences,
State University of New York,
Albany, NY 12222, USA

Thomas J. Valone
Department of Biology,
Concordia University,
1455 ouest, Boulevard de Maisonneuve,
Montréal, Québec
H3G 1M8, Canada

Animals often acquire foraging skills through learning (e.g., Croy and Hughes, 1990; Sullivan, 1988). We may dichotomize learning as either "individual" or "social," although the two processes often interact in social groups (Galef, 1988). For our purposes, individual learning involves only a direct transaction between a consumer and its resource. The process of individual learning is the same for a solitary and a group member; social influences do not directly affect the individual's interaction with the resource and subsequent acquisition of a novel behavior. However, the probability that a group member encounters the opportunity to learn individually may depend on behavioral attributes of other group members (Giraldeau, 1984).

Social learning, for our purposes, envisions direct effects of individuals performing a resource-producing behavior on an observer's acquisition of that behavior. In general, social learning may imply that group members acquire innovative behavior faster than solitaries for two reasons. First, knowledgeable individuals may channel an observer's attention toward stimuli or conditions promoting acquisition of the innovation. Examples include social facilitation (contagion of motivational states) and local enhancement (directing attention to particular locations or objects; Thorpe, 1956: 120–124). Second, an observer may acquire an innovation by imitating knowledgeable "tutors." Experimental evidence of imitation of novel food-finding skills has been provided for pigeons (e.g., Palameta and Lefebvre, 1985) and for rats (e.g., Heyes and Dawson, 1990). The probability that an individual group member learns an innovation through either of these social effects should depend on the frequency of the behavior among other group members (e.g., Bartholomew, 1983; Boyd and Richerson, 1985, 1988).

Social learning, in its various forms, drives cultural evolution both within and between generations (Fagan, 1981; Lefebvre, 1986; Lefebvre and Palameta, 1988; Pulliam and Dunford, 1980). Our models examine within-generation, frequency-dependent effects on both individual learning and cultural transmission of a resource-producing skill in groups of social foragers. It is often the case in such groups that some individuals (scroungers) obtain food discovered/captured by other group members (producers). Our theory demonstrates how the producer-scrounger phenomenon (e.g., Barnard and Sibly, 1981; Caraco and Giraldeau, 1991; Parker, 1984; Vickery et al., 1991) can interact with learning processes to govern behavioral diversity within a social group.

Recent experimental studies of learning have shown that social foraging can interfere with individual learning (Beauchamp and Kacelnik, 1991; Fragaszy and Visalberghi, 1989; Giraldeau and Lefebvre, 1986). Scrounging interferes with individual learning because scroungers (who would otherwise be exploring the environment and learning to produce resources) spend so much time eating at food clumps discovered by other group members (Giraldeau, 1984). Furthermore, scrounging may occasionally interfere with social learning (Fragaszy and Visalberghi, 1990; Giraldeau and Lefebvre, 1987; Giraldeau and Templeton, 1991). This interference may have ecological relevance because empirical evidence suggests that the rate at which individuals acquire or improve the efficiency of foraging skills can have important fitness consequences (Croy and Hughes, 1990; Sullivan, 1988).

A few previous models examine the simultaneous effects of both individual learning and within-generation cultural transmission on the frequency of individuals exhibiting a particular trait (Bartholomew, 1983; Boyd and Richerson, 1985). However,

T. J. Valone is now at the Department of Biology, University of New Mexico, Albuquerque, NM 87131, USA.

Received 19 March 1992
Revised 9 September 1992
Second Revision 5 November 1992
Accepted 13 November 1992
1045-2249/94/$5.00

the complications of scrounging have not been previously considered in modeling the dynamics of a learned resource-producing trait. Our first model assumes that scrounging inhibits any social learning, as suggested by Giraldeau and Lefebvre (1987). Consequently, the model investigates how factors such as scrounging, foraging group size, and skill complexity might affect the equilibrium frequency of an innovation acquired only through individual learning. A second model relaxes the effect of scrounging on social learning (see Giraldeau and Templeton, 1991) and assumes that both individual and social learning may occur. The results lead to new predictions concerning the reductions in individual learning caused by scrounging. The models suggest that the ecological significance of social learning may lie partially in its capacity to circumvent the inhibitory effects that scrounging has on foraging skill acquisition through individual learning.

MODELS

Our theory concerns the learned capacity to produce resource clumps that are shared by members of a closed social group. We assume that both individual and social learning of a trait enhancing resource production depend on that trait's frequency within the group. Frequency dependence of social learning makes intuitive sense; an increased number of "demonstrators" exhibiting an economically advantageous trait implies an increase in the opportunities for others to learn socially (e.g., Bartholomew, 1983; Cavalli-Sforza and Feldman, 1981; Pulliam and Dunford, 1980). Therefore, the probability that a given individual acquires the trait socially may increase with the trait's frequency. Frequency dependence for individual learning may be less obvious, but no less important. We assume that as the frequency of the resource-producing trait increases, a naive individual has fewer opportunities to learn the trait as a result of its own experience discovering resources. That is, increasing the likelihood a naive individual will share (scrounge) a resource clump discovered by another forager decreases the chance the first individual will acquire the trait through individual learning (Giraldeau, 1984).

Each of our models considers two levels of resource production. Individuals discover resource clumps at either a baseline rate or a faster rate; the faster rate of clump discovery is the trait acquired through learning. The models admit the possibility of forgetting and relearning the trait and examine the equilibrium frequency distribution of the trait within a foraging group. We use numerical analysis to determine the effects that group size, task complexity, and consequence of learning on resource production rate have on the equilibrium numbers of individuals that have and have not acquired the skill.

Individual learning only: long-term equilibrium

A foraging group composed of G members searches for food resources in a patchy environment. When one individual discovers (or otherwise produces) a food clump, each group member consumes (scrounges) a portion of the resource. At any given time we can expect some variation among group members in the rate of discovering clumps (e.g., Caraco et al., 1989). To focus the model, we assume that each group member can be classified as either a slow or fast producer. Slow producers discover clumps at a constant, lower probabilistic rate, s_1, and fast producers discover clumps at a greater, constant probabilistic rate, s_2, where:

$$s_2 = \alpha s_1; \qquad \alpha > 1. \tag{1}$$

A slow producer may increase its rate of clump discovery through individual learning only. That is, a forager's learning requires that the individual itself discover resource clumps; having discovered a clump, a slow producer may learn from the experience and become a fast producer. Because $s_2/s_1 = \alpha$, a larger value of α implies a greater consequence of learning on producing rate and a larger increase in each group member's feeding rate when learning occurs. A large α may occur when, for instance, the spontaneous rate of performance of some behavior is very low, but when the behavior is performed it leads to the discovery of an abundant food supply.

We also assume that forgetting can occur, although forgetting may be rare compared to learning. When an individual forgets, by definition, a fast producer becomes a slow producer, so that its rate of clump discovery reverts from s_2 to s_1. Sociological analyses of the intragenerational dynamics of learned traits usually allow forgetting and relearning (e.g., Karmeshu and Pathria, 1979; Sharma et al., 1982, 1983). More importantly, empirical results imply that forgetting occurs in both mammalian (D'Amato, 1973) and avian foragers (Giraldeau and Lefebvre, 1987; Grant and Roberts, 1973) when opportunities to use a learned trait are rare or when tasks are complex (Commons, 1981).

As the frequency of group members that learned the trait increases, the probability that the next clump is discovered by any given slow producer declines. Consequently, a slow producer's opportunities for learning also decline as the number of fast producers increases. The negative frequency dependence of individual learning in our model affects the equilibrium proportion of individuals that exhibit the trait as we vary group size, the relative likelihood of forgetting versus learning, and the relative resource discovery rates of fast and slow producers.

At time t, a group of G foragers contains $X(t)$ slow producers and $[G - X(t)]$ fast producers: $X(t) \in \{0, 1, \ldots, G\}$. Individuals discover resource clumps independently at respective probabilistic rates s_1 and s_2. At time t, the probability that the next clump is discovered by any of the slow producers is $\theta(t)$:

$$\begin{aligned}\theta(t) &= s_1\, X(t)/\{s_1\, X(t) + s_2\, [G - X(t)]\} \\ &= X(t)/\{X(t) + \alpha[G - X(t)]\}.\end{aligned} \tag{2}$$

For $X(t) < G$, the likelihood that any of the slow producers learns to discover resources at the higher rate will be proportional to $\theta(t)$.

We assume $X(t)$ is a stochastic process. The set $\{X(t_1), X(t_2), \ldots, X(t_z)\}$ has the same probability distribution as $\{X(t_1 + \tau), X(t_2 + \tau), \ldots, Xt_z + \tau)\}$, so the process is stationary (see Kelly, 1979). We assume $X(t)$ has the Markov property; the present

Table 1

Individual learning only for group size (G) of 5 ($Y = G - X$)

		α		
	(μ/β)	2	5	10
$E[X]$	0.01	0.4	0.78	1.21
	0.1	2.09	2.7	3.12
	1.0	4.23	4.33	4.39
$V[X]$	0.01	0.36	0.59	0.75
	0.1	1.09	0.91	0.72
	1.0	0.62	0.47	0.39
$CV[X]$	0.01	1.49	0.98	0.72
	0.1	0.5	0.35	0.27
	1.0	0.19	0.16	0.14
$CV[Y]$	0.01	0.13	0.18	0.23
	0.1	0.36	0.41	0.45
	1.0	1.02	1.02	1.02

state contains all useful information for predicting future states. The only allowable transitions in $X(t)$ are increases by 1 (for $X(t) < G$) and decreases by 1 (for $X(t) > 0$). Therefore, $X(t)$ constitutes a birth–death process on a finite-state space.

Suppose $X(t)$ increases from r to $(r + 1)$. Then one of $(G - r)$ fast producers forgets the learned trait. We assume that any forgetting occurs independently among the fast producers. The probabilistic transition rate from r to $(r + 1)$ is $q(r, r + 1)$:

$$q(r, r + 1) = \mu(G - r); \qquad r = 0, 1, \ldots, G - 1, \tag{3}$$

where μ is the individual rate of forgetting. The value of μ may be relatively small, but it should increase as either task complexity or the rarity of clump discovery increases.

Now suppose $X(t)$ decreases from r to $(r - 1)$. Then one of the slow producers learns to discover resources at the higher rate. The probabilistic transition rate from r to $(r - 1)$ is $q(r, r - 1)$. Using Equation 2:

$$q(r, r - 1) = \beta r/\{r + \alpha(G - r)\}; \qquad r = 1, 2, \ldots, G, \tag{4}$$

where β is the learning parameter; greater values of β should be associated with less complex tasks. Note that the transition rate per naive individual, $q(r, r - 1)/r$, increases as r increases. Hence a slow producer's probability of learning individually decreases as the trait's frequency ($[G - r]/G$) increases.

Because $X(t)$ is a stationary birth–death process, its properties ensure that it will have a unique equilibrium probability distribution (Kelly, 1979). We designate the equilibrium distribution $p(X)$ where:

$$p(x) = \lim_{t \to \infty} Pr[X(t) = X \mid X(0) = k], \tag{5}$$

The proportion of time the process spends in state X converges to $p(X)$ as $t \to \infty$, independently of the initial state of the process $X(0)$.

Then the equilibrium probability function, as derived in Appendix A, is:

$$p(X) = \frac{\binom{G}{X}[\mu(\alpha - 1)/\beta]^{x}[\Gamma(C)/\Gamma(C - X)]}{\sum_{i=0}^{G}\binom{G}{i}[\mu(\alpha - 1)/\beta]^{i}[\Gamma(C)/\Gamma(C - i)]} \tag{6}$$

for $X = 0, 1, \ldots, G$; $C = \alpha G/(\alpha - 1)$ and $\Gamma(k) = (k - 1)\Gamma(k - 1)$. Equation 6 defines the equilibrium distribution of X, the number of animals that have not yet learned the skill. The ratio (μ/β) increases with task complexity and may decrease with environmental productivity (i.e., clump density). The quantity α increases as the consequence of learning on an individual's rate of clump discovery increases. Of course, no forgetting ($\mu = 0$) means that every forager acquires and retains the trait, so that $p(0) = 1$ at equilibrium.

We numerically analyzed $p(X)$, the equilibrium number of individuals that have not learned, across a range of values for group size (G), task complexity (μ/β), and consequences of learning on foraging rate (α). For each parameter combination, we calculated the equilibrium number that have not learned and then found the mean ($E[X]$), variance ($V[X]$), and coefficient of variation ($CV[X]$) of the equilibrium distribution. As the mean of the equilibrium number that fail to learn increases, the mean of the number that learn decreases. An increase in the variance of the number of those that have not learned is mirrored in the variance of the number that have learned:

$$V[G - X] = V[G] + V[X] - 2\,COV[G,X]. \tag{7}$$

But G is a constant, so $V[G - X] = V[X]$.

The coefficient of variation in the number of those that have not learned is $\sqrt{V[X]}/E[X]$. The equivalent quantity for the number that have learned is:

$$\begin{aligned} CV[G - X] &= \sqrt{V[G - X]}/E[G - X] \\ &= \sqrt{V[X]}/(G - E[X]). \end{aligned} \tag{8}$$

An increase in $CV[X]$ may imply an increase or a decrease in $CV[G - X]$ for each $p(X)$.

Table 1 lists $E[X]$, $V[X]$, $CV[X]$ and $CV[G - X]$ for $G = 5$ at nine different combinations of task complexity (μ/β) and α, the consequences of learning on producing rate. Table 2 shows parallel re-

Table 2

Individual learning only for group size (G) of 10 ($Y = G - X$)

		α		
	(μ/β)	2	5	10
$E[X]$	0.01	1.5	2.66	3.74
	0.1	5.82	6.71	7.41
	1.0	9.15	9.24	9.32
$V[X]$	0.01	1.2	1.63	1.76
	0.1	1.92	1.76	1.22
	1.0	0.74	0.59	0.42
$CV[X]$	0.01	0.73	0.48	0.35
	0.1	0.24	0.2	0.15
	1.0	0.09	0.08	0.07
$CV[Y]$	0.01	0.13	0.17	0.21
	0.1	0.33	0.4	0.43
	1.0	1.01	1.01	0.96

Table 3
Individual learning only for group size (G) of 15 ($Y = G - X$)

		α		
	(μ/β)	2	5	10
$E[X]$	0.01	3.1	5.11	6.79
	0.1	9.94	11.12	11.91
	1.0	14.11	14.19	14.27
$V[X]$	0.01	2.27	2.69	2.67
	0.1	2.89	2.09	1.59
	1.0	0.8	0.67	0.56
$CV[X]$	0.01	0.49	0.32	0.24
	0.1	0.17	0.13	0.11
	1.0	0.06	0.06	0.05
$CV[Y]$	0.01	0.19	0.17	0.2
	0.1	0.34	0.37	0.41
	1.0	1.0	1.01	1.03

sults for $G = 10$, and Table 3 shows results for $G = 15$. Examination of the tables reveals patterns within and among levels of group size.

For a given group size, the expected equilibrium number that have not learned increases as task complexity and as the consequence of learning on producing rate increase. This result follows directly from the assumption that the probability a slow producer discovers a clump (and perhaps learns the task) decreases as the consequence of learning on producing rate increases. Neither of these results is surprising.

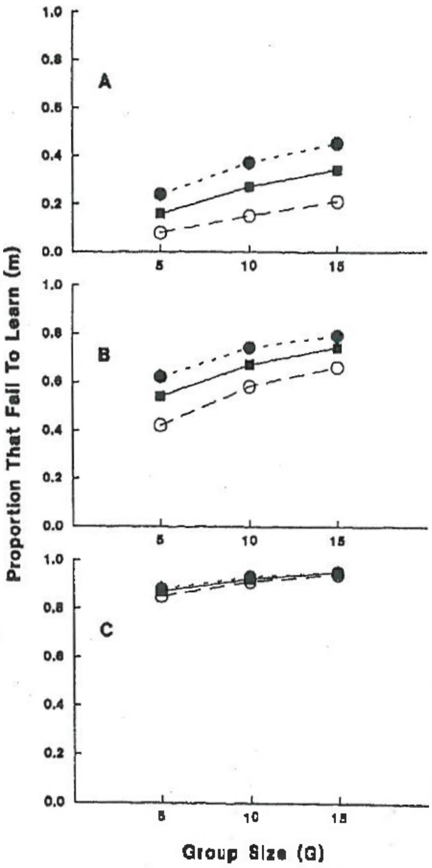

Figure 1
The proportion of group members (m) that fail to learn a food-finding skill as a function of group size (G), when behavioral complexity of the skill is (A) low ($\mu/\beta = 0.01$), (B) intermediate ($\mu/\beta = 0.1$), and (C) high ($\mu/\beta = 1.0$). Dashed line, $\alpha = 2$; solid line, $\alpha = 5$; dotted line, $\alpha = 10$.

Note that the expected number of individuals that have not acquired the skill in the group exceeds the number expected if each individual were foraging as a solitary individual. When each of G individuals forages solitarily (hence independently), the average number of slow producers at equilibrium is $G\mu/(\beta + \mu) = G(\mu/\beta)/(1 + [\mu/\beta])$. For given μ and β, this quantity is less than $E[X]$ in our model (equivalently, the number of individuals having acquired the skill is smaller in our model) because the opportunity for individual learning exhibits frequency dependence in a social group.

When group size is fixed, our model's variance in the number of individuals that have not acquired the skill becomes smaller when the mean number that have not acquired the skill approaches extremes of either 0 or G, and the variance is larger when the mean is intermediate, near $G/2$. Consequently, the variance of the number that have not acquired the skill does not vary monotonically with increases in either skill complexity or consequence of learning on producing rate. Because the mean number that have not learned always increases with either skill complexity or consequences of learning on producing rate, patterns in the coefficients of variation of individuals that have and have not learned depend more on variation in the mean than on variation in the variance.

For a given group size, the coefficient of variation in the expected number of individuals that have not learned declines as either skill complexity or consequence of learning on producing rate increases. Because these patterns reflect increases in the mean number that have not learned, the coefficient of variation of the number that have learned increases as either skill complexity or consequence of learning on producing rate increases. The coefficient of variation in the number of group members failing to exhibit the learned trait should decrease as either task complexity or the consequence of learning on producing rate increases. The coefficient of variation in the number exhibiting the trait should vary in the opposite direction.

The above considerations held group size fixed, while other parameters were varied. An interesting pattern emerges when we examine coefficients of variation across group size. As group size increases, the coefficient of variation of the number that have not learned clearly decreases with all parameter combinations. However, the coefficient of variation of the number that have learned remains roughly constant.

At equilibrium, the average proportion of individuals in the group that have not yet learned is simply $m = E[X]/G$. Figure 1 shows this proportion plotted as a function of group size. The figure indicates that the average proportion of individuals failing to perform the learned trait should increase as task complexity, consequence of learning on producing rate, or group size increases. Consequently, the expected proportion of the learned task declines in larger groups.

The preceding model restricts learning to an outcome of individual experience in producing a resource. The model's results include the following predictions: (1) the proportion of individuals exhibiting a learned trait should decrease as either task complexity (μ/β) or the consequence of learning on producing rate (α) increases; (2) the proportion of individuals exhibiting a learned trait should decrease as group size increases; and (3) the coefficient of variation of the number of individuals failing to exhibit the trait should decline with increasing group size, but the coefficient of variation of the number exhibiting the trait should be independent of group size.

If no forgetting occurs, each individual eventually learns the same resource-producing trait. Our dichotomous characterization of phenotypes says little about within-group diversity unless individuals sometimes forget the task. Suppose we eliminated forgetting in the above model by letting $\mu = 0$. For this case, instead of analyzing the number of individuals that have not learned at equilibrium, we take the time elapsing until $X(t) = 0$, i.e., the time until all group members have learned the trait, as an index of phenotypic diversity. A longer duration of the learning process promotes the chance of observing within-group variation.

At time $t = 0$, we set $X(t) = G$, so that no group member exhibits the trait for fast resource production. The variable T represents the total time for all group members to learn the trait. (In Appendix B we derive the approximate mean and variance of T.) The approximate expected value of T is:

$$E[T] \approx (G/\beta)\ [1 - \alpha + \alpha \ln(G+1)]. \tag{9}$$

The approximate variance of T is:

$$V[T] = (G/\beta^2)\left\{\alpha\left[\frac{\pi^2}{6} - \frac{1}{G+1}\right] - 2(\alpha - 1)[1 + \alpha \ln(G + 1)]\right\}. \tag{10}$$

Both the mean and variance of the total time increase as group size increases and decreases as task complexity decreases (β increases). The mean always increases as the consequence of learning on producing rate increases for $G \geq 2$, but the variance of T declines as the consequence of learning on producing rate increases. These results are consistent with the predictions concerning the frequency of the learned trait in the model with forgetting.

Individual and social learning: long-term equilibrium

The first model showed how an increased frequency of the trait for greater resource production should impede individual learning in naive individuals. However, an increased number of fast producers could provide additional opportunities for social transmission of the trait through social facilitation, local and stimulus enhancement, or imitation, even though scrounging mave have some interfering effect on social learning (see Giraldeau and Templeton, 1991). The following model opposes the effects that a trait's frequency has on individual learning with the effect it has on its cultural transmission.

As above, $X(t)$ is a stationary, stochastic process counting the number of group members not exhibiting the learned trait at time t, and $s_2 = \alpha\, s_1$, as in the preceding model. If $X(t)$ increases from r to $(r + 1)$, a fast producer forgets the production skill. The probabilistic transition rate, $q(r, r + 1)$, again is $\mu(G - r)$ for $r = 0, 1, \ldots, G - 1$. If $X(t)$ decreases from r to $(r - 1)$, a slow producer learns to discover resource clumps at the higher rate. We assume individual learning proceeds exactly as in the first model; the probabilistic rate of transition due to individual learning is $q^*(r, r - 1) = \beta r/[r + \alpha(G - r)]$ for $r = 1, 2, \ldots, G$.

We make a standard assumption about cultural transmission: probabilistic transition rates will be proportional to the product of the numbers of slow and fast producers. That is, the total rate of social learning should be proportional to the number of demonstrators multiplied by the number of individuals that might learn (e.g., Bartholomew, 1983; Cavalli-Sforza and Feldman, 1981; but see Lefebvre and Giraldeau, in press). So the probabilistic rate of transition due to social learning is $\hat{q}(r, r - 1)$, where:

$$\hat{q}(r, r - 1) = \Omega r(G - r); \qquad r = 1, 2, \ldots, G, \tag{11}$$

and Ω is the contagion parameter representing the combined effectiveness of the fast producers' demonstrations and the slow producers' social learning ability. The total transition rate due to social learning increases with r, the number of naive individuals, until that number exceeds one-half the group size, after which the total rate declines as the number of naive individuals continues to increase. The transition rate of social learning per naive individual, $\hat{q}(r, r - 1)/r$, decreases as r increases. Hence the probability of cultural transmission per naive individual increases as the frequency of the trait $[(G - r)/G]$ increases, opposite to the frequency-dependent effect assumed for individual learning.

Assuming that individual and social learning are probabilistically independent, the total transition rate from r to $(r - 1)$ is the sum of q^* and $\hat{q}$:

$$q(r, r - 1) = r\left[\frac{\beta}{r + \alpha(G - r)} + \Omega(G - r)\right], \tag{12}$$

for $r = 1, 2, \ldots, G$. Equations 3 and 12 define a birth–death process with an equilibrium distribution $p(X)$, satisfying Equations A2a and A2b.

In terms of the transition rates, $p(X)$ is:

$$p(X) = p(0) \prod_{r=1}^{X} [\mu(G - r + 1)/r] \cdot \left[\frac{\beta}{r + \alpha(G - r)} + \Omega(G - r)\right]^{-1}$$

$$p(X) = \binom{G}{X}\mu^X p(0) \cdot \prod_{r=1}^{X} \frac{r(1 - \alpha) + \alpha G}{\alpha\Omega(G - r)^2 + \Omega r(G - r) + \beta}.$$

We can simplify the multiplicand as in Appendix A and use Equation A2b to find the expression for $p(0)$. For simplicity we define $M(X)$ as:

$$M(X) = \prod_{r=1}^{X} \alpha\Omega(G - r)^2 + \Omega r(G - r) + \beta.$$

Then the equilibrium probability function is:

Table 4
Individual and social learning for group size (G) of 5 and β of 0.6 ($Y - G - X$)

		α	
	Ω	2	10
$\mu = 0.2$			
$E[X]$	0.3	0.83	0.89
	0.9	0.28	0.28
$V[X]$	0.3	0.88	0.97
	0.9	0.29	0.29
$CV[X]$	0.3	1.13	1.11
	0.9	1.92	1.91
$CV[Y]$	0.3	0.22	0.24
	0.9	0.11	0.11
$\mu = 1.2$			
$E[X]$	0.3	3.99	4.1
	0.9	2.24	2.32
$V[X]$	0.3	1.44	1.28
	0.9	2.74	2.8
$CV[X]$	0.3	0.3	0.28
	0.9	0.74	0.72
$CV[Y]$	0.3	1.19	1.26
	0.9	0.6	0.62

$$p(X) = \frac{\text{[illegible]}\,[\mu(\alpha - 1)]^X[\Gamma(C)/\Gamma(C - X)M(X)]}{\sum_{i=0}^{G}\binom{G}{i}[\mu(\alpha - 1)^i[\Gamma(C)/\Gamma(C - i)M(i)]}, \tag{13}$$

for $X = 0, 1, \ldots, G$; $C = \alpha G/(\alpha - 1)$.

We again analyzed numerically the equilibrium number of individuals who have not learned. We fixed the value of task complexity and varied group size, forgetting rate, consequences of learning on producing rate, and the contagiousness of the skill.

Table 5
Individual and social learning for group size (G) of 10 and β of 0.6 ($Y = G - X$)

		α	
	Ω	2	10
$\mu = 0.2$			
$E[X]$	0.3	0.74	0.73
	0.9	0.25	0.25
$V[X]$	0.3	0.75	0.69
	0.9	0.25	0.25
$CV[X]$	0.3	1.17	1.15
	0.9	2.02	2.01
$CV[Y]$	0.3	0.01	0.01
	0.9	0.01	0.01
$\mu = 1.2$			
$E[X]$	0.3	6.26	5.35
	0.9	1.51	1.58
$V[X]$	0.3	3.43	6.72
	0.9	1.57	1.87
$CV[X]$	0.3	0.3	0.48
	0.9	0.83	0.87
$CV[Y]$	0.3	0.5	0.48
	0.9	0.15	0.16

Tables 4 and 5 present results for groups of 5 and 10 individuals, respectively. Because the number of parameters in Equation 13 is relatively large, we might assemble a long list of interactive effects. Instead, we focus on patterns that examine the generality of the results obtained from the previous model.

For a given group size, the expected number of individuals that have not acquired the skill decreases as the contagion parameter increases. Greater task complexity, whether interpreted as increased (μ/β) or a decreased contagion parameter (Ω), results in greater numbers of individuals that have not learned. But the average number of group members failing to perform the task is less sensitive to variation in the consequence of learning on producing rate in comparison to the first model. Generally, the consequence of learning on producing rate is less influential when both individual and social learning occur. When we considered individual learning only, the frequency of the fast-producing trait was lower in a group of size G than among G independent, solitary foragers. Social learning can overcome this effect. A sufficiently large contagion parameter increases the proportion of individuals acquiring the skill (equivalently, decreases $E[X]$) beyond the level expected among solitaries that must learn individually and independently.

With group size held constant, the coefficient of variation of the number of individuals that have not learned increases as the contagion parameter increases; this pattern reflects the inverse relationship between $E[X]$ and Ω. The coefficient of variation of the number of individuals that have learned correspondingly decreases as the contagion parameter increases, unless the proportion of individuals having acquired the skill is close to unity (see Table 5 where $\mu = 0.2$).

Next we consider patterns between group sizes. In contrast to the first model, increasing group size sometimes results in a reduced expected number of individuals that do not learn. Social learning may be more likely to overcome the frequency-dependent restraint on individual learning when groups are larger (see below).

As group size increases, the coefficient of variation of the number of individuals that have not learned generally increases or remains essentially unchanged. This contrasts to the previous model where the coefficient of variation in the number that have not learned and group size were inversely related. The coefficient of variation in the number that have learned declines with group size; this quantity was independent of group size in the first model. The expected proportion of group members not performing the task (m) decreases as group size increases. That is, the learned trait will be exhibited by a greater proportion of individuals as group size increases when both types of learning occur. This result contrasts to the corresponding result in the first model.

In summary, with individual learning only, the frequency of the learned trait decreased with both task complexity and the consequence of learning on producing rate and decreased as group size increased. With both individual and social learning, task complexity had the same effect: the consequence of learning on producing rate failed to affect the proportion of individuals having acquired the skill, and the expected proportion of individuals using the skill increased as group size increased. With individual learning only, the coefficient of variation of the number not performing the task declined with group size, and the coefficient of variation of the number performing the trait was independent of group size. With both individual and social learning, the first coefficient of variation was roughly independent of group size, and the second declined with group size. Frequency dependence of learning processes within social groups requires close attention to the relative strengths of individual and social learning before any generalities concerning their effects on phenotypic diversity can be attempted.

DISCUSSION

Our models develop the consequences of allowing the frequency of a learned trait to govern an individual's opportunities to acquire the trait through both individual and social learning. We focus on the way scrounging influences the spread of foraging innovations within groups. If the spatio-temporal distribution of food allows scrounging, individual learning is an increasingly inefficient mechanism of skill acquisition as group size (G), skill complexity μ/β), or the effect of learning on performance (α) increases. Social learning can overcome this inefficiency, especially in larger groups. After describing our model's relationship to other analyses of social learning, we discuss our results in light of the possible consequences for understanding the evolution of cultural transmission.

In general, cultural transmission may occur within or between generations. Empirical documentation of the diffusion of traits significantly affecting foraging success suggests that the within-generation models are the more appropriate device (Lefebvre and Palameta, 1988). Our models consider only a single group. A logical extension might examine a population stratified into separate groups where individual learning and relatively strong cultural transmission occur within groups, and weaker cultural transmission occurs between groups (e.g., Coleman, 1964). Presumably, the number of resource-producing traits that might be learned, the degree of stratification in the dynamics of social learning, and frequency dependence of learning opportunities will interact with the economic consequences of possessing particular traits to influence the overall trait diversity among individuals.

Stephens (1991) scales temporal variability in the environment to generation length and suggests conditions favoring the evolution of learning over a fixed mapping of genotype on phenotype. Little attention has been directed to questions of why animals that possess individual learning mechanisms should also evolve the capacity for social learning. It does appear that individual and social learning are not simply manifestations of a single "domain-general" mechanism of cognition; they are probably different special-purpose adaptive mechanisms (Tooby and Cosmides, 1989). Consequently, individual and social learning can be treated as distinct, but perhaps dependent, attributes (Boyd and Richerson, 1988; Lefebvre and Palameta, 1988). Of course, in many situations the two will interact; social learning may initiate the acquisition of a skill that is later honed through individual experience.

Most ideas concerning the evolution of social learning simply assume that it is faster or less costly than individual learning (e.g., Boyd and Richerson, 1985; Galef, 1976). But hypothesizing an overall superior efficiency for social learning lacks an appreciation of the remarkably varied ecological circumstances where learned traits may enhance survival and reproduction. Lefebvre and Palameta (1988) take an ecological perspective on the evolution of social learning. They suggest that an opportunistic, highly diversified diet should promote reliance on social learning, as copying others allows the individual to exploit new food types as they become available (Galef, 1988).

Boyd and Richerson (1988) develop a model of environmental tracking to assess conditions favoring between-generation social learning. They assume that increased social learning implies a decrease in individual learning and that individual learning is a more costly mechanism. They conclude that social learning (essentially, mimicry) should either be more widespread than is currently believed or that it is more costly than often assumed (Boyd and Richerson, 1988).

Social learning probably is more rapid than individual learning, but it must often impose costs in terms of time budgeting. Imitators will need to spend time observing tutors and will need to invest some time discriminating between relevant and irrelevant behavioral sequences of tutors. Time spent watching tutors could detract from antipredator vigilance or from searching for resources the individual does know how to exploit. Therefore, social learning is not necessarily less costly than individual learning.

The relative costs of individual and social learning may be unimportant if the benefit of acquiring a particular resource-producing skill is large. Our models assume no differential cost between the two learning modes. We also assume that the probability of forgetting is independent of the learning mechanism, but the mechanism of learning governs the likelihood of forgetting. Little empirical information is currently available concerning that possibility.

Our models show that when scrounging is possible, individual learning becomes increasingly inefficient for skill acquisition. Moreover, the inefficiency is greater for those skills that lead to larger increases in food discovery rates. Even if social learning is as costly as individual learning, scrounging can promote the evolution of social learning as a way of avoiding this inefficiency. It follows that ecological conditions increasing the occurrence of scrounging, such as larger group size and greater spatio-temporal clumping of resources (Caraco and Giraldeau, 1991; Vickery et al., 1991), also can drive the evolution of social learning and hence culture.

APPENDIX A

Equilibrium distribution: individual learning only

We want to identify $p(X)$, the equilibrium distribution of the number of group members failing to exhibit the learned trait. Because $X(t)$ is a birth–death process with the Markov property, the equilibrium distribution must satisfy the "detailed" balance conditions (Kelly, 1979):

$$p(r)/p(r-1) = q(r-1, r)/q(r, r-1). \quad \text{(A1)}$$

Starting with $r = 1$, one can derive a general formula for the equilibrium distribution from the balance conditions:

$$p(X) = p(0) \prod_{r=1}^{X} [q(r-1, r)/q(r, r-1)] \quad \text{(A2a)}$$

for $X = 1, 2, \ldots, G$; and

$$p(0) = 1 - \sum_{X=1}^{G} p(X). \quad \text{(A2b)}$$

Discussion of birth–death processes as models for social organization can be found in Boswell et al. (1979), Caraco (1980), and Cohen (1972).

Using Equations 3 and 4 from the text, we have $q(r-1, r) = \mu(G - r + 1)$, and $q(r, r-1) = \beta r/[r + \alpha(G - r)]$. Then Equation A2a becomes:

$$p(X) = p(0) \prod_{r=1}^{X} \mu(G - r + 1) \cdot r + \alpha(G - r)]/\beta r. \quad \text{(A3)}$$

Simplification yields:

$$p(X) = \binom{G}{X} (\mu/\beta)^X p(0) \prod_{r=1}^{X} [r + \alpha(G - r)].$$

The multiplicand can be written as a ratio of gamma functions (Caraco, 1979), and substituting the last expression into Equation A2b gives an expression for $p(0)$.

APPENDIX B

Mean and variance of the time until each group member acquires the trait

With individual learning and no forgetting, the only allowable transitions decrease $X(t)$ from r to $(r - 1)$. The associated transition rates $q(r, r-1)$ are given by Equation 4. We let the random variable t_r represent the duration of the process in the state $X(t) = r$. Each t_r has an exponential probability density with mean $E[t_r] = [q(r, r-1)]^{-1}$, and variance $V[t_r = [q(r, r-1)]^{-2}$. The variable T represents the total time for all group members to learn the trait and t is a sum of random variables:

$$T = \sum_{r=1}^{G} t_r. \quad \text{(B1)}$$

The expected value of T is:

$$E[T] = \sum_{r=1}^{G} [q(r, r-1)^{-1}$$

$$= \sum_{r=1}^{G} [r + \alpha(G - r)]/\beta r. \quad \text{(B2)}$$

This expression becomes:

$$E[T] = (G/\beta)(1 - \alpha) + (\alpha G/\beta) \sum_{r=1}^{G} (1/r). \quad \text{(B3)}$$

Approximating the partial sum of the harmonic series yields:

$$E[T] \approx (G/\beta)[1 - \alpha + \alpha \ln(G + 1)]. \quad \text{(B4)}$$

The variance of T is $V[T]$:

$$V[T] = \sum_{r=1}^{G} [q(r, r-1)]^{-2}$$

$$= \sum_{r=1}^{G} [r + \alpha(G - r)]^2/(\beta r)^2. \quad \text{(B5)}$$

After expanding, we obtain:

$$V[T] = (1 - \alpha)(2G/\beta^2) + (1 - \alpha)(2\alpha G/\beta^2) \sum_{r=1}^{G} (1/r) + (\alpha G/\beta)^2 \sum_{r=1}^{G} (1/r^2). \quad \text{(B6)}$$

Applying standard approximations for the partial sum yields:

$$V[T] \approx (G/\beta^2)\left\{\alpha\left[\frac{\pi^2}{6} - \frac{1}{G+1}\right] - 2(\alpha - 1)[1 + \alpha \ln(G + 1)]\right\}. \quad \text{(B7)}$$

Financial support for this research was provided by the Natural Sciences and Engineering Research Council (Canada) University Research Fellowship and Operating Grant, a Fonds pour la Formation de Chercheurs et L'Aide à la Recherche (Québec) Nouveaux Chercheurs Grant (to L.-A.G.), and a Hasselblad Foundation grant (to T.C.). We thank M. Björklund, A. Kacelnik, K. Laland, and K. Nishimura for useful comments.

REFERENCES

Barnard CJ, Sibly RM, 1981. Producers and scroungers: a general model and its application to feeding flocks of house sparrows. Anim Behav 29:543–550.

Bartholomew DJ, 1983. Stochastic models for social processes. New York: Wiley.

Beauchamp G, Kacelnik A, 1991. Effects of the knowledge of partners on the learning rates of zebra finches *Tainiopygia guttata*. Anim Behav 41:247–253.

Boswell MT, Ord JK, Patil GP, 1979. Chance mechanisms underlying univariate distributions. In: Statistical distributions in ecological work (Ord JK, Patil JP, Taillie C, eds). Burtonsville, Maryland: International Co-operative Publishing; 3–156.

Boyd R, Richerson PJ, 1985. Culture and the evolutionary process. Chicago: University of Chicago Press.

Boyd R, Richerson PJ, 1988. An evolutionary model of social learning: the effects of spatial and temporal variation. In: Social learning: psychological and biological perspectives (Zentall TR, Galef BG Jr, eds). Hillsdale, New Jersey: Lawrence Erlbaum Associates; 29–48.

Caraco T, 1979. Ecological response of animal group size frequencies. In: Statistical distributions in ecological work (Ord JK, Patil GP, Taillie C, eds). Burtonsville, Maryland: International Co-operative Publishing; 371–386.

Caraco T, 1980. Stochastic dynamics of avian foraging flocks. Am Nat 115:262–275.

Caraco T, Barkan C, Beacham JL, Brisbin L, Lima S, Mohan A, Newman JA, Webb W, Withiam ML, 1989. Dominance and social foraging: a laboratory study. Anim Behav 34:41–58.

Caraco T, Giraldeau L-A, 1991. Social foraging: producing and scrounging in a stochastic environment. J Theor Biol 153:559–583.

Cavalli-Sforza LL, Feldman MW, 1981. Cultural transmission and evolution: a quantitative approach. Princeton, New Jersey: Princeton University Press.

Cohen, JE, 1972. Markov population processes as models for primate social and population dynamics. Theor Popul Biol 3:119–134.

Coleman JS. 1964. Introduction to mathematical sociology. London: Collier-Macmillan.

Commons ML, 1981. How reinforcement density is discriminated and scaled. In: Quantitative analyses of behavior, vol. 1: discriminative properties of reinforcement schedules (Commons ML, Nevin JA, eds). Cambridge, Massachusetts: Ballinger; 51–85.

Croy MI, Hughes RN, 1990. Combined effects of learning and hunger in the feeding behaviour of the fifteen-spined stickleback (*Spinachia spinachia*). In: Behavioural mechanisms of food selection (Hughes RN, ed). Berlin: Springer-Verlag; 215–234.

D'Amato MR, 1973. Delayed matching and short-term memory in monkeys. In: The psychology of learning and motivation: advances in research and theory (Bower GH, ed). New York: Academic Press; 27–48.

Fagen RM, 1981. Animal play behavior. New York: Oxford University Press.

Fragaszy DM, Visalberghi E, 1989. Social influences on the acquisition of tool-using behaviours in tufted capuchin monkeys (*Cebus apella*). J Comp Psychol 103: 159–170.

Fragaszy DM, Visalberghi E, 1990. Social processes affecting the appearance of innovative behaviors in capuchin monkeys. Folia Primatol 54:155–165.

Galef BG Jr, 1976. Social transmission of acquired behavior: a discussion of tradition and social learning in vertebrates. In: Advances in the study of behavior, vol. 6 (Rosenblatt JS, Hinde RA, Shaw E, Beer C, eds). New York: Academic Press; 77–100.

Galef BG Jr, 1988. Imitation in animals: history, definition, and interpretation of data from the psychological laboratory. In: Social learning: psychological and biological perspectives (Zentall TR, Galef BG Jr, eds). Hillsdale, New Jersey: Lawrence Erlbaum Associates; 3–28.

Giraldeau L-A, 1984. Group foraging: the skill pool effect and frequency-dependent learning. Am Nat 124:72–79.

Giraldeau L-A, Lefebvre L, 1986. Exchangeable producer-scrounger roles in a captive flock of feral pigeons: a case for the skill pool effect. Anim Behav 34:797–803.

Giraldeau L-A, Lefebvre L, 1987. Scrounging prevents cultural transmission of food-finding behaviour in pigeons. Anim Behav 35:387–394.

Giraldeau L-A, Templeton JJ, 1991. Food scrounging and diffusion of foraging skills in pigeons, *Columba livia*: the importance of tutor and observer rewards. Ethology 89:63–72.

Grant DS, Roberts WA, 1973. Trace interaction in pigeon short-term memory. J Exp Psychol 101:21–29.

Heyes CM, Dawson GR, 1990. A demonstration of observational learning in rats using a bidirectional control. Q J Exp Psychol 42B:59–71.

Karmeshu, Pathria RK, 1979. Cooperative behaviour in a non-linear model of diffusion of information. Can J Phys 57:1572–1578.

Kelly FP, 1979. Reversibility and stochastic networks. New York: Columbia University Press.

Lefebvre L, 1986. Cultural diffusion of a novel food-finding behaviour in urban pigeons: an experimental field test. Ethology 71:295–304.

Lefebvre L, Giraldeau L-A, in press. Cultural transmission in pigeons is affected by the number of tutors and bystanders present during demonstrations. Anim Behav.

Lefebvre L, Palameta B, 1988. Mechanisms, ecology, and population diffusion of socially learned, food-finding behavior in feral pigeons. In: Social learning: psychological and biological perspectives (Zentall TR, Galef BG Jr, eds). Hillsdale, New Jersey: Lawrence Erlbaum Associates; 141–164.

Palameta B, Lefebvre L, 1985. the social transmission of a food-finding technique in pigeons: what is learned? Anim Behav 33:892–896.

Parker GA, 1984. The producer/scrounger model and its relevance to sexuality. In: Producers and scroungers: strategies of exploitation and parasitism (Barnard CJ, ed). Beckenham, England: Croom Helm; 127–153.

Pulliam HR, Dunford C, 1980. Programmed to learn: an essay on the evolution of culture. New York: Columbia University Press.

Sharma CL, Pathria RK, Karmeshu, 1982. Critical behavior of a class of nonlinear stochastic models of diffusion of information. Phys Rev A 26:3567–3674.

Sharma CL, Pathria RK, Karmeshu, 1983. Diffusion of information in a social group. J Math Sociol 9:211–226.

Stephens DW, 1991. Change, regularity, and value in the evolution of animal learning. Behav Evol 2:77–89.

Sullivan KA, 1988. Ontogeny of time budgets in yellow-eyed juncos: adaptation to ecological constraints. Ecology 69:118–124.

Thorpe, WH, 1956. Learning and instincts. London: Methuen.

Tooby J, Cosmides L, 1989. Evolutionary psychology and the generation of culture, part I. Ethol Sociobiol 10: 29–49.

Vickery WL, Giraldeau L-A, Templeton JJ, Kramer DL, Chapman CA, 1991. Producers, scroungers and group foraging. Am Nat 137:847–863.