A Theory of Associating Food Types with Their Postingestive Consequences

Jonathan M. Yearsley,[1,*] **Juan J. Villalba,**[2,†] **Iain J. Gordon,**[3,‡] **Ilias Kyriazakis,**[4,§] **John R. Speakman,**[5,‖]
Bert J. Tolkamp,[4,#] **Andrew W. Illius,**[6,**] **and Alan J. Duncan**[1,††]

1. Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH, United Kingdom;
2. Department of Forest, Range, and Wildlife Services, Utah State University, Logan, Utah 84322;
3. Sustainable Ecosystems Division, Commonwealth Scientific and Industrial Research Organization, Davies Laboratory, Townsville, Queensland 4810, Australia;
4. Animal Nutrition and Health Department, Scottish Agricultural College, Edinburgh EH9 3JG, United Kingdom;
5. Rowett Research Institute, Aberdeen AB21 9SB, United Kingdom;
6. School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JR, United Kingdom

ABSTRACT: Animals often face complex and changing food environments. While such environments are challenging, an animal should make an association between a food type and its properties (such as the presence of a nutrient or toxin). We use information theory concepts, such as mutual information, to establish a theory for the development of these associations. In this theory, associations are assumed to maximize the mutual information between foods and their consequences. We show that associations are invariably imperfect. An association's accuracy increases with the length of a feeding session and the relative frequency of a food type but decreases as time delay between consumption and postingestive consequence increases. Surprisingly, the accuracy of an association is independent of the number of additional food types in the environment. The rate of information transfer between novel foods and a forager depends on the forager's diet. In light of this theory, an animal's diet may have two competing goals: first, the provision of an appropriate balance of nutrients, and second, the ability to quickly and accurately learn the properties of novel foods. We discuss the ecological and behavioral implications of making associational errors and contrast the timescale and mechanisms of our theory with those of existing theory.

Keywords: diet selection, postingestive feedback, mutual information.

* Corresponding author. Present address: Département d'Écologie et Évolution, Université de Lausanne, CH-1015 Lausanne, Switzerland; e-mail: jonathan.yearsley@unil.ch.

† E-mail: villalba@cc.usu.edu.

‡ E-mail: iain.gordon@csiro.au.

§ E-mail: ilias.kyriazakis@sac.ac.uk.

‖ E-mail: j.speakman@rri.sari.ac.uk.

E-mail: bert.tolkamp@sac.ac.uk.

** E-mail: a.illius@ed.ac.uk.

†† E-mail: a.duncan@macaulay.ac.uk.

Diet choice underpins the ecological relationships between an animal and its food resources. Ecological theory interprets this diet choice as an adaptive behavior, where selective forces shape the food types that an animal includes in its diet in order to maximize its fitness (Stephens and Krebs 1986; Houston and McNamara 1999). The extensive literature on the adaptive theory of diet selection makes one common assumption, that an animal associates a food type with its fitness consequences (e.g., energy density, nutrient content, toxin concentration; Schoener 1971; Pulliam 1974; Westoby 1974; Stephens and Krebs 1986; Farnsworth and Illius 1998). For this association to be made, the food must inform the animal about its properties, the animal must make a choice between foods that differ in these properties, and the animal must relate these properties to their fitness consequences. Although some theory has looked at the value of foraging information (Stephens 1989), theories have not considered how this information is established and whether it is likely to be maintained in a world where both foods (van Marken Lichtenbelt 1993; Bautista et al. 1998; Provenza et al. 2003) and animal states (Houston 1993; Newman et al. 1995; Kyriazakis et al. 1999; Villalba and Provenza 1999) vary over time.

Experimental studies across a range of animal species (e.g., ruminants [Provenza 1995; Behmer et al. 2002], rats [Garcia and Kimeldorf 1960; Garcia and Hankins 1977], chickens [Covasa and Forbes 1994], coyotes and wolves [Gustavson et al. 1976], and cephalopods [Gelperin 1975;

Susswein et al. 1986; Darmaillacq et al. 2004]) show that animals can associate a food type with its postingestive consequence and develop a diet preference in accordance with a beneficial or detrimental postingestive consequence. This ability is remarkable because the conditioned stimulus (e.g., the taste of a food) and the unconditioned stimulus (the postingestive consequence) can be separated by many minutes or hours (Garcia et al. 1966; Westoby 1974; Forbes 1999; Arsenos et al. 2000). In contrast, the conditioned and unconditioned stimuli for most other learned associations occur within seconds of one another (e.g., Pavlov 1927; Carlson 1994). This ability to learn from a delayed reinforcement is important because a temporal delay always exists between the orosensory experience and the postingestive consequence of a food (Provenza 1995; Sclafani 1995).

The observed ability of animals to learn from postingestive consequences has become a cornerstone in our understanding of the mechanisms of diet selection (Provenza 1995). However, experiments that demonstrate conditioned food aversion and preference generally establish the aversion with a single food type (e.g., Provenza et al. 1994; Villalba et al. 1999). This is very different from the feeding environment faced by wild animals, which live in a world with a complex mix of different food types. Free-ranging animals can consume a variety of food types in one feeding session, which requires simultaneously developing associations between several food types and their postingestive consequences. It is suggested that several foods, all with different nutrient concentrations, ingested in short succession will decrease the probability of correct association (van Wieren 1996). This view is supported by experimental evidence that demonstrates that herbivores can be compromised in their ability to develop associations when several food types are available simultaneously (Duncan and Young 2002). It seems that an animal's ability to establish an accurate association between a food type and its postingestive consequences is limited and depends on the feeding environment and perhaps the animal's feeding behavior (e.g., the frequency of inclusion of new foods in the diet). If diet selection in the real world presents the animal with too many novel foods for it to make reliable associations, then any extrapolation from simple experimental conditions that deal with one or two stimuli may be unwarranted.

In this article, we develop a theory that describes how food properties and foraging behavior can affect the accuracy of an association between a food type and its postingestive consequence. These associations have ecological and behavioral implications because they describe how much of an animal's diet is founded not on the true properties of a food but on an erroneous association with other properties in the feeding environment. The theory predicts the important factors governing these associations (e.g., food and environmental properties and foraging behavior), which can in turn suggest mechanisms of prey aggregation, frequency-dependent selection, foraging behavior, and the evolution of prey defense mechanisms. Our theoretical framework also suggests a new hypothesis for an animal's diet selection when confronted with novel food types for which the postingestive consequences are unknown. We emphasize that our study ignores the long-term fitness implications of the postingestive consequences so that it can isolate the important processes for making an association. Inclusion of long-term fitness consequences is not expected to change the process of association and is discussed later on. Our theory, therefore, radically differs from classical theories of diet selection because it is wholly concerned with the accuracy of food type associations rather than the fitness consequences of a food type. Both processes are expected to play a role in explaining diet because both accurate association of food types and their fitness consequences are important aspects of a forager's diet selection.

Methods

Specifying the Problem

We start by analyzing a diet that contains two food types and then generalize these results to n food types. The problem of associating n food types with their respective postingestive consequences, $c_1, \ldots, c_n$, can be formulated in terms of an association between two time series (fig. 1); one time series is the list of food types that have been ingested, and the other time series is the list of postingestive consequences (these consequences can be beneficial or detrimental to an animal). The time series are divided into feeding sessions. Each session is one time step in
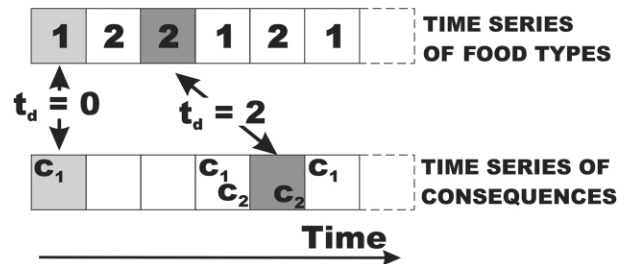
Figure 1: Representation of the situation faced by an animal that has to associate two food types with their consequences. Food types 1 and 2 are associated with consequences c_1 and c_2, respectively. Consequence c_1 has no delay in its effects ($t_d = 0$), while consequence c_2 is delayed from the time of ingestion by $t_d = 2$. This time delay complicates the association between the time series of food types and their consequences.

length, and during each session only one food type is consumed. The action of not eating can be incorporated into this formalism as an additional "phantom" food type with its own consequence (the consequence of not eating is to experience no postingestive signal). If the postingestive consequences are all experienced very soon after ingestion, then it is straightforward to associate them with each food type. However, if postingestive consequences are delayed, as is often the case (Provenza 1995; Sclafani 1995), then establishing associations is more complicated because a time delay requires a memory of the past and increases the likelihood of interactions with other sensory and postingestive signals. Our formulation of the problem has two important assumptions: first, given a food time series, an animal's diet selection attempts to maximize the amount of information about the postingestive consequences; and second, older information is remembered with decreasing resolution.

The first assumption amounts to a behavior where the distinction between food types (i.e., the difference between their estimated consequences) is maximized. This distinction is equivalent to the information transfer between the time series of food types and the time series of consequences and can be quantified by drawing on concepts developed in information theory (Weaver and Shannon 1949; Wells 1999). The transfer of information between two time series is formally defined by their mutual information, *I*. Mutual information is the answer to the question, "Given a single measurement from one process, how much information on average can this measurement predict about a second process?" The mutual information between one time series of *n* food types and another of *n* postingestive consequences is defined as

$$I \equiv \sum_{i=1}^{n} Q_i I(i) \tag{1}$$

$$I(i) \equiv \sum_{j=1}^{n} \pi_{ij} \log_2 \left(\frac{\pi_{ij}}{P_j} \right), \tag{2}$$

where $I(i)$ is the partial mutual information of consequence c_i, π_{ij} is an animal's perceived probability that food type j is associated with postingestive consequence c_i (π_{ij} is also known as the forward transmission probability), P_j is the relative frequency of food type j in the food time series, and Q_i is the relative frequency of postingestive consequence c_i. The mutual information has a minimum $I = 0$, corresponding to complete uncertainty ($\pi_{ij} = P_{ij}$), and a maximum of $\log_2(n)$, corresponding to the maximum information content of the food time series. In all practical applications, the maximum mutual information will not be attained. A related application of mutual in-

formation has been suggested in the context of finding time delays in time series data (Fraser and Swinney 1986).

As feeding information is recalled from further in the past, the uncertainty of the information increases (Weber's law would suggest that the uncertainty is proportional to the time delay [Bateson and Kacelnik 1998], although recent evidence suggests that time perception has a nonlinear component [Crystal 2001]). This uncertainty can be realized in many different forms. One possible form is shown in figure 2, where the estimated time delay of a consequence is $\hat{t}_d$, and all food types from the present back to $\hat{t}_d$ are included as possible sources of the present consequence. The system in figure 2 associates a time delay with each consequence. An alternative and equivalent representation would be to associate the time delays with the different food types. The important mechanism is that there is an uncertainty in associating a food type with its consequence, and this uncertainty becomes increasingly important as the delay in the postingestive consequence increases.

While this approach assumes that a forager can make use of past information, it is difficult to relate this to a forager's cognitive ability because the use of past information does not require a high level of cognition. Our approach could apply to a range of foragers with varying cognitive abilities.

Describing the Food Time Series

For the purposes of this article, the food time series is described as a Markov chain (Grimmett and Stirzaker
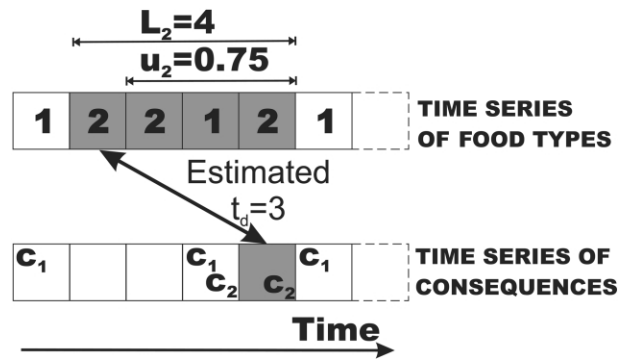
Figure 2: One possible model for the uncertainty involved in trying to associate a food type with its postingestive consequence. In this example, the consequence c_2 is estimated to have a time delay of $\hat{t}_d = 3$. Given this time delay estimate, all food types consumed for three time units before the present are considered as possible sources of the present consequence (shaded region in the food time series). Using this time delay estimate, an animal would estimate that c_2 has a 75% probability of being associated with food type 2. This corresponds to $L_2 = t_d + 1 = 4$ and $u_2 = 0.75$.

1992). In general this need not be the case, but a Markov chain enables us to explicitly calculate the probabilities required in equations (1) and (2) while incorporating autocorrelation into the time series (it is a first-order autoregressive process). The probability that an animal switches food types and starts feeding on a new food type depends on its preferences for each food type and on the food type on which it is currently feeding. Increasing autocorrelation within the time series corresponds to longer series of feeding sessions on any one food type.

Two Food Types.   We start by developing the Markov chain model for two food types, labeled 1 and 2, that have post-ingestive consequences c_1 and c_2 (figs. 1, 2). The probability that an animal's present feeding session is on food type 1 is given by P_1 (if an animal has no preference for any food type, then P_1 is the frequency of food type 1 in the environment). Given that an animal is eating food type 1, the probability that it feeds on the same food type in the next feeding session is α_1. This conditional parameter defines the amount of serial autocorrelation in the animal's feeding time series for food type 1. If $\alpha_1 = 0.5$, then feeding is random; if $\alpha_1 > 0.5$, then feeding on food type 1 is clustered; and if $\alpha_1 < 0.5$, then feeding on food type 1 is dispersed. The parameters P_1 and α_1 can be used to generate a time series of food types, and the properties of this time series can be used to calculate expected values for π_{ij} in equation (1).

Of particular interest to our analysis are the probabilities that food type 1 and food type 2 occur in the time series, given that food type 1 occurred t time steps ago (represented as $p_{11}[t]$ and $p_{21}[t]$, respectively), because these probabilities are required in the calculation of mutual information (eq. [1]). These probabilities can be calculated as

$$p_{11}(t) = 1 - p_{21}(t) = P_1 + (1 - P_1)\Delta^t, \qquad (3)$$

where

$$\Delta = \alpha_1 + \alpha_2 - 1 \qquad (4)$$

describes the overall autocorrelation in the time series for both food types. The autocorrelation parameter ranges from strong negative autocorrelation when $\Delta = -1$ ($\alpha_1 = \alpha_2 = 0$) to strong positive autocorrelation when $\Delta = 1$ ($\alpha_1 = \alpha_2 = 1$). The probabilities $p_{11}(t)$ and $p_{21}(t)$ lie in the range from 0 to 1. Putting these ranges into equation (3) then shows that for certain values of Δ, the value of P_1 cannot be arbitrarily small. This is a consequence of the Markov chain model because transitions from food type 2 to food type 1 impose a minimum frequency for P_1.

Given that food type 2 occurred t time steps ago, the equivalent of equation (3) is

$$p_{22}(t) = 1 - p_{12}(t) = P_2 + (1 - P_2)\Delta^t. \qquad (5)$$

The covariance matrix for the presence of food type 1 (and food type 2), given that food type i occurred t and t' time steps ago (where $t' \geq t$), can also be calculated as

$$\sigma_{1i}(t, t') = \sigma_{2i}(t, t') = p_{1i}(t)p_{2i}(t)\Delta^{t'-t}. \qquad (6)$$

The variance in occurrence of food type 1, given that food type i occurred t time steps ago, is $\sigma_{1i}(t, t)$.

Many Food Types.   The model for two food types described above can be extended to look at n food types. The Markov chain for n food types can be described by

$$x_i(t + 1) = \alpha_i x_i(t) + \sum_{j \neq i} \frac{s_i}{1 - s_j}(1 - \alpha_j)x_j(t), \qquad (7)$$

where $x_i(t)$ is the probability that an animal feeds on food type i at time t, and s_i is the selective preference for food type i, such that $\sum_{i=1}^{n} s_i = 1$. Equation (7) gives the probability of occurrence of food type i as $P_i \propto (1 - s_i)/(1 - \alpha_i)$. More complicated models for n food types are possible, and these would allow for more general transition probabilities between food types, but these details go beyond the scope of this article.

Focusing on food type 1, we can rewrite equation (7) as a two-food-type Markov chain:

$$x_1(t + 1) = \alpha_1 x_1(t) + [1 - \alpha_0(t)]x_0(t), \qquad (8a)$$

$$x_0(t + 1) = \alpha_0(t)x_0(t) + (1 - \alpha_1)x_1(t), \qquad (8b)$$

where $x_0(t)$ is the probability that an animal does not feed on food type 1 at time t, and $\alpha_0(t)$ is the conditional probability that an animal continues not to feed on food type 1. The variables $x_0(t)$ and $\alpha_0(t)$ can be expressed as

$$x_0(t) \equiv \sum_{i \neq 1} x_i(t) = 1 - x_1(t), \qquad (9a)$$

$$\alpha_0(t) = 1 - \sum_{i \neq 1} s_1 \frac{x_i(t)}{x_0(t)} \frac{1 - \alpha_i}{1 - s_i}. \qquad (9b)$$

The equilibrium properties of the Markov chain for n food types can then be calculated by substituting P_i for $x_i(t)$ in equation (9). Equation (9b) also allows the autocorrelation parameter Δ (eq. [4]) to be rewritten from the perspective of food type 1 as

$$\Delta = \alpha_1 + \alpha_0(t) - 1. \qquad (10)$$

Equations (9) and (10) allow the original two-food-type model to be extended so that time series with n food types can be generated with specific values for P_i and the autocorrelation parameter Δ (see below). Furthermore, if $\alpha_0(t)$ is time independent, then all the analytical results from the two-food model can be applied to models with n food types. This will be the case when $(1 - \alpha_j)/(1 - s_j)$ is constant for all but food type 1 (i.e., P_i is constant for all but food type 1). In general, $\alpha_0(t)$ is time dependent, and the relationship between the time evolution of the n-food-type model and the two-food-type model is more complicated. For this more general case we have not found analytical solutions analogous to equations (3)–(6).

Simulating Food Time Series. Feeding time series are generated using equation (7) by defining P_i, s_i, and Δ. These parameters specify all the conditional probabilities α_i and allow food type frequency and autocorrelation to be controlled (for two food types, s_i is not specified because there is no choice when selecting a new food type). Increasing the number of food types above two requires that the frequencies and selective preferences of the additional food types be adjusted to account for the changing number of food types, while P_1 and Δ are kept constant.

We have two approaches for specifying food type frequency as a function of the number of food types. For both approaches, the frequency of food type 1 is always constant. In the first approach, the frequency of the remaining food types are $P_j = (1 - P_1)/(n - 1)$. In this approach, $\alpha_0(t)$ is time invariant, and all the analytical results for two food types can be applied. In the second approach, P_j is allocated using a broken-stick model. So the unallocated portion of the time series is split in half (other ratios can be used), giving $P_j = (1 - P_1)0.5^{j-1}$, except for the last food type, whose frequency is $P_n = (1 - P_1)0.5^{n-2}$.

We also look at two approaches of specifying s_i. In the first approach, s_i is a constant for all food types. This assumes that an animal picks one of the possible $n - 1$ food types at random. Food type preferences are then reflected through the length of time an animal spends feeding on a food type. In the second approach, we let $s_i = P_i$, so that an animal selects a food type according to its frequency in the food time series. In this approach, food type preferences are reflected through the initial selection of the foods as well as the length of time an animal spends feeding on a food type. These two approaches for assigning s_i are considered only when frequencies are assigned using the broken-stick approach above because the results are not affected by s_i when frequencies are equal, $P_j = (1 - P_1)/(1 - n)$.

Associating Postingestive Consequences with Food Types

The probability π_{ij} of making an association between a food type and a postingestive consequence is required for the calculation of mutual information (eqq. [1], [2]). There are a number of ways of interpreting the probability of a correct association, π_{ii}. The simplest considers associations as being either perfect or imperfect, so that in a population of animals, π_{ii} is the proportion of animals that make a perfect association of food type i with consequence c_i. A second interpretation allows associations to be partially correct, so that an animal can consider a food type's consequence to be a mixture of the consequences being experienced. In this interpretation, π_{ii} is the component of the estimated consequence that is correct. For all interpretations, the probability of making an incorrect association for food type i is $1 - \pi_{ii}$. So π_{ij} (where $i \neq j$) is the probability that an animal incorrectly associates food type i with consequence c_j.

To estimate π_{ij}, consider an animal that is experiencing the postingestive consequence c_1. The animal associates this postingestive consequence with a section of the feeding time series, which has a length L_1. There are several plausible mechanisms underlying such an association. For example, the animal may have some estimate of the time delay on c_1, in which case L_1 would represent the uncertainty in remembering the food intake time series after this time delay (memory in this case need not be conscious because conditioned food aversions can be subconscious). Figure 2 shows this scenario, where the uncertainty covers all food types from the present back to the beginning of the estimated time delay (in this case, $t_d = 3$).

Given the model shown in figure 2, mutual information can be estimated directly from the data by a range of methods (e.g., Paninski 2003). However, using equations (3)–(6), we can also derive analytical equations for the mutual information in our model. We consider the case where the true feeding event giving rise to c_1 is captured within the section L_1 (it can be shown that mutual information always decreases if this is not the case). Based on the presence of the postingestive consequence c_1 and L_1, the expected probability that food type 1 is associated with c_1 can be calculated from equation (3) as

$$\pi_{11} = \sum_{t=0}^{(L_1-1)u_1} \frac{p_{11}(t)}{L_1} + \sum_{t=0}^{(L_1-1)(1-u_1)} \frac{p_{11}(t)}{L_1}, \qquad (11)$$

which can be rewritten using equation (3) as

$$\pi_{11} = 1 - \pi_{21} = P_1 + (1 - P_1)A(L_1, u_1), \qquad (12a)$$

where $A(L_1, u_1)$ is given by

$$A(L_i, u_i) = \frac{1}{L_i}\left\{1 + \frac{\Delta}{(1 - \Delta)}\left[2 - \Delta^{u_i(L_i-1)} - \Delta^{(L_i-1)(1-u_i)}\right]\right\}$$

(12*b*)

and u_i is defined as the proportional distance along the time series section where the true feeding event associated with c_i occurs. The quantity $A(L_i, u_i)$ takes a value between 0 and 1 and describes the deviation from randomness in the association of consequence c_i. If $A(L_i, u_i) = 0$, then the association is random (i.e., $\pi_{ii} = P_i$), while if $A(L_i, u_i) = 1$, then $\pi_{ii} = 1$ and there is complete certainty in associating c_i with food type i. The equivalent equation for c_2 can be obtained from equation (5) as

$$\pi_{22} = 1 - \pi_{12} = P_2 + (1 - P_2)A(L_2, u_2). \quad (12c)$$

The mutual information of an association between food types and postingestive consequences can now be calculated by substituting equation (12) into equation (2). Figure 3 shows an example of the mutual information for two food types over a range of estimated time delays. In figure 3, the animal's best estimates of the time delays for the two food types occur at the points of maximum mutual information (i.e., $\hat{t}_{d1} = 26$ and $\hat{t}_{d2} = 13$).

The variance of the mutual information can also be calculated from the variance about the expectations in equation (12). Using equations (6) and (11), we can find the variance in π_{ij} as

$$\mathrm{Var}(\pi_{1i}) = \mathrm{Var}(\pi_{2i})$$

$$= \frac{V_i[(L_i - 1)u_i] + V_i[(L_i - 1)(1 - u_i)]}{L_i^2}, \quad (13a)$$

where

$$V_i(s) = \sum_{t,t'=0}^{s} \sigma_{ii}(t, t'), \quad (13b)$$

from which the variance in the mutual information can be calculated as

$$\mathrm{Var}(I) = \sum_{i=1}^{2} 2Q_i^2 \left(\log_2 \frac{\pi_{i1}P_2}{\pi_{i2}P_1}\right) \mathrm{Var}(\pi_{ii}). \quad (14)$$

The analysis can be simplified if the partial mutual information of a food type (eq. [2]) is independent of the associations of the other postingestive consequences (i.e., if the values of $[L_i, u_i]$ for each consequence are independent of each other). If this is the case, then the maximum partial mutual information for consequence c_i is given by
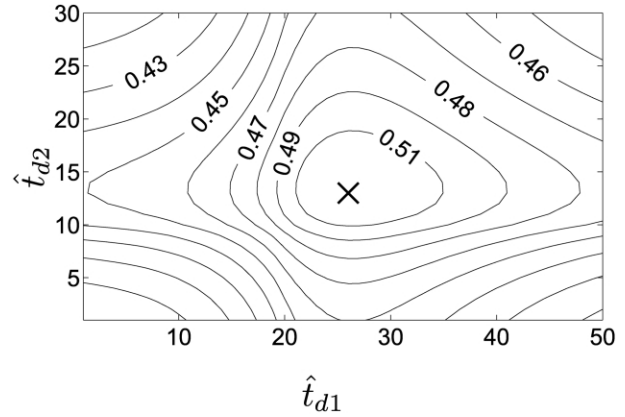
Figure 3: Mutual information in associating two food types with their postingestive consequences (eq. [1]) for a range of estimated time delays, $\hat{t}_{d1}$ and $\hat{t}_{d2}$. The frequency of food type 1 is $P_1 = 0.7$; the time delays for c_1 and c_2 are $t_{d1} = 20$ and $t_{d2} = 10$, respectively. The maximum mutual information occurs at time delay estimates of $\hat{t}_{d1} = 26$ and $\hat{t}_{d2} = 13$ for c_1 and c_2, respectively (*cross*). The autocorrelation in the feeding time series is $\Delta = 0.9$.

the value of π_{ii} that is furthest from random (randomness is $\pi_{ii} = 1/n$, where n is the number of food types). This simplifies the analysis because equation (12) need not be substituted into equation (2).

For any one postingestive consequence, the parameters L_i and u_i are not independent because both are determined by the estimated time delay for the consequence, $\hat{t}_{di}$. For example, in figure 2 the length of a portion of time series is given by the estimated time delay, making $L_1 = \hat{t}_{d1} + 1$ and $u_1 = (\hat{t}_{d1} - t_{d1})/L_1$, where t_{d1} and $\hat{t}_{d1}$ are the actual and estimated time delays for c_1. Other relationships could be constructed; for example, L_1 could be proportional to the estimated time delay and centered on this time delay.

Using this formalism, we now investigate how the probability of a correct association depends on the time delay of a food's consequence, the relative frequency of a food type, the number of food types in the diet, and the length of a feeding session (i.e., the autocorrelation).

Results

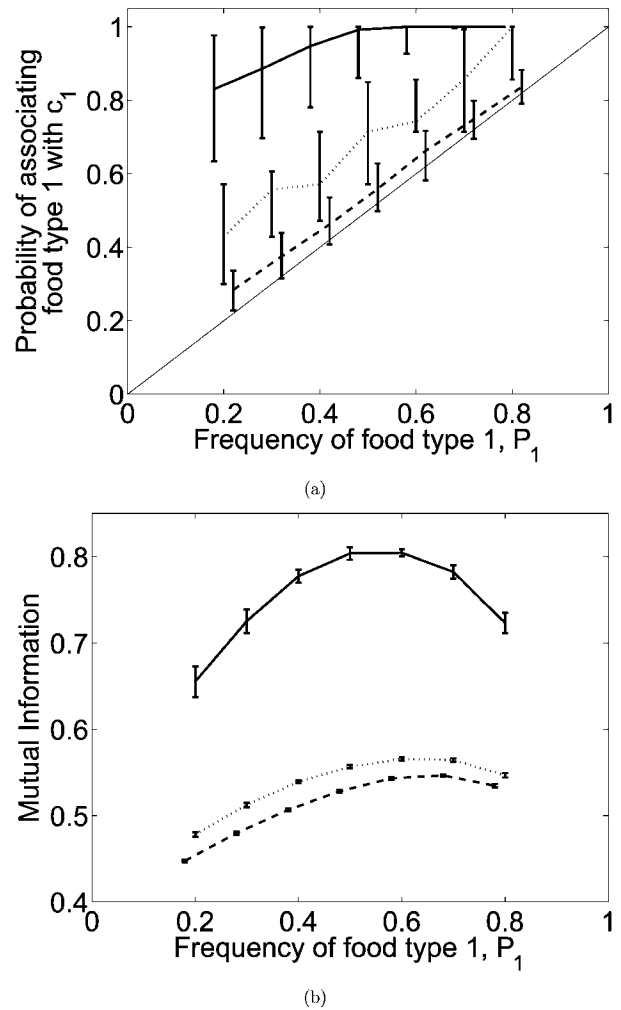### Relative Frequency of Food Types

The relative frequency of a food type in the feeding time series is the only parameter that causes an interaction between different food types because an increase in the frequency of one food type requires a decrease in the frequency of another. The probability of a correct association between a food type and its consequence generally decreases with the relative frequency of a food type. However, the relative frequency of a food type has a threshold value

that separates two behaviors of the median probability of association (fig. 4*a*). Above the threshold frequency, the median probability of correct association is close to unity and weakly dependent on frequency, while the confidence intervals around the median decrease with increasing frequency. In contrast, below the threshold frequency, the median probability of association decreases with frequency (associated with decreases of $A[L_i, u_i]$ in eq. [12]), while the confidence intervals are relatively independent of frequency. This threshold frequency decreases as the autocorrelation in the feeding time series increases or the time delay associated with a postingestive consequence decreases (fig. 4*a*). The effect of autocorrelation is understood by considering a fixed length of time series. Increasing autocorrelation decreases the number of food types in this length of time series, which increases the probability of a correct association (i.e., makes associations less random). Reducing time delay also reduces the number of food types, but it does this by reducing length of time series used to make an estimate. The interaction between the median and its confidence intervals (fig. 4*a*) gives a mean probability of an association that is a linear function of food type frequency, with no threshold behavior (eq. [12]).

The amount of transmitted information is best quantified by the mutual information (fig. 4*b*). The maximum mutual information is a well-known quantity in communication theory known as the capacity of the communication channel. The food type frequency that achieves this capacity depends on the time delays associated with all postingestive consequences and the autocorrelation in the feeding time series. In figure 4*b*, the maximum mutual information occurs when food type 1 has a frequency of $P_1 \approx 0.6$ in an animal's feeding time series. If this frequency is increased, information about food type 2 is lost, while if this frequency is decreased, information about food type 1 is lost. The capacity of the communication channel at $P_1 \approx 0.6$ provides a balance between these two processes where the flow of total information about the whole system is maximized. The mutual information can also be partitioned into the partial information for each food type (eq. 2), which quantifies the information transfer about any one food type.

Number of Food Types

When $\alpha_0(t)$ (eq. [9*b*]) is time independent, all the results of the two-food-type model can be applied to the many-food-type scenario. The important conclusion from this is that the accuracy of an association is independent of the number of additional food types in an environment. When $\alpha_0(t)$ is time dependent, we use simulations to study the effect of food type number. These simulations show

(a)

(b)

Figure 4: Best-estimated probability of making a correct association (*a*) and mutual information (*b*) as functions of the frequency of food type 1 in the feeding time series P_1 (calculated from 20 time series, each 10,000 feeding sessions long). The faint solid line without error bars in *a* shows a random association (i.e., probability of association is equal to frequency of occurrence). Solid and dashed lines denote postingestive consequences with a time delay of $t_{d1} = 20$, with $\Delta = 0.95$ and $\Delta = 0.5$, respectively, while the dotted line denotes $t_{d1} = 5$ and $\Delta = 0.5$. The second food type has a postingestive consequence with $t_{d2} = 1$ and identical autocorrelation. Error bars represent 35% and 65% quantiles.

that if the relative frequency of a food type is constant, then the probability of making a correct association is remarkably independent of the number of food types in an environment. This was found to be the case irrespective of whether s_i was dependent on the relative frequency of a food type.

The number of food types can affect the accuracy of an association if changing the number of food types changes the relative frequency of a food type. Increasing the num-
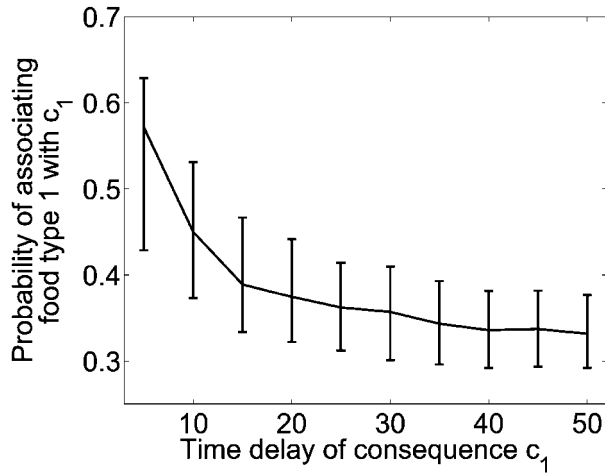
Figure 5: Best-estimated probability of making a correct association as a function of the time delay of a postingestive consequence (calculated from 20 time series, each 10,000 feeding sessions long). The association is calculated from the best estimate of the time delay of c_1. The frequency of food type 1 is $P_1 = 0.3$, its autocorrelation in the feeding time series is $\Delta = 0.5$, and the time delay of food type 2 is $t_{d2} = 10$ with identical autocorrelation. Error bars represent 35% and 65% quantiles.

ber of food types may generally be expected to decrease the relative frequency of a food type, which in turn will reduce the accuracy of an association. However, this is primarily an effect of relative frequency rather than of the number of food types.

Time Delay of a Postingestive Effect

The optimum time delay estimate for c_1 can be seen from figure 3 to be $\hat{t}_{d1} \approx 26$, which is longer than the actual time delay selected for this scenario ($t_{d1} = 20$) because the feeding time series is positively autocorrelated ($\Delta = 0.9$). Positive autocorrelation increases the time delay estimate because a food type and its postingestive consequence are more likely to occur in clusters. This means that a more accurate signal of the true consequences can be gained by including feeding events on either side of the feeding session that gave rise to the present effect.

Figure 5 shows the effect of the time delay for c_1 on the median probability of making a correct association. This gives the intuitively reasonable result that an increasing time delay produces a monotonic decline in the expected probability of making a correct association. For large time delays, the probability of making a correct association approaches the relative frequency of the food type (i.e., a random association). Increasing the time delay of consequence c_i increases the length of the time series that is used to estimate an association, and this can only make

associations more random. The uncertainty around the median probability of correct association tends to be highest when the probability of association is close to 0.5 (eqq. [6], [13]).

Autocorrelation

Increasing positive autocorrelation in the feeding series (increasing Δ) increases the clustering of food types, which results in longer periods of feeding on the same food type. Differentiating equation (12b) with respect to Δ shows that increasing the autocorrelation in a time series always makes an association less random, decreasing the uncertainty in an association (fig. 6). Figure 6 also shows that increasing the autocorrelation widens the confidence intervals around the median (the same can be seen for the variance from eqq. [6] and [14]).

The effect of the autocorrelation also depends on the time delay because the time delay affects both L_i and u_i. Figure 6 shows that as the time delay of an effect increases, the impact of autocorrelation decreases. In this sense the time delay and the autocorrelation are opposing effects.

Discussion

Classical diet selection theory (Stephens and Krebs 1986) proposes that an animal's choice of diet is determined by the properties of the food types in its environment (e.g.,
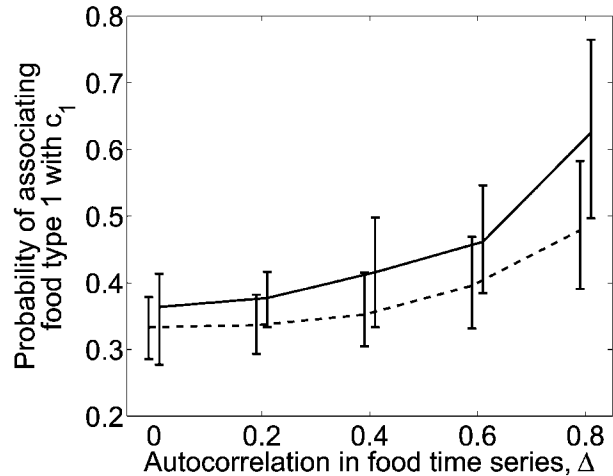

Figure 6: Best-estimated probability of making a correct association as a function of the autocorrelation in the feeding time series (calculated from 20 time series, each 10,000 feeding sessions long). Solid and dashed lines denote postingestive effects with time delays of $t_{d1} = 10$ and $t_{d1} = 20$, respectively. The frequency of food type 1 is $P_1 = 0.3$, and the time delay for consequence of food type 2 is $t_{d2} = 10$. Error bars represent 35% and 65% quantiles.

handling time, energy, and toxin content). Extending this concept, we propose that an animal's diet selection is also motivated by its ability to accurately associate these properties with the appropriate food type. Incorrect associations between food types and their properties will weaken or possibly reverse the strength of evolutionary selection. We use mutual information to quantify this association. The time delays for each postingestive consequence are estimated by finding the strongest association (i.e., the maximum mutual information), and from this the probability that a food type is associated with a postingestive consequence can be estimated. The results produce the following testable hypotheses: first, decreasing the time delay between food ingestion and postingestive consequence increases the accuracy of an association; second, decreasing the number of consecutive foraging sessions on one food type reduces the accuracy of an association; third, decreasing the relative abundance of a food type decreases the accuracy of correctly associating it with its consequences and increases the probability that the food type is associated with the consequences from other food types in the environment; and finally, increasing the number of food types in a forager's diet has a very weak effect on the accuracy of the association of a food type whose relative abundance in the diet is held constant. The theory also shows that the rate of learning new food types (i.e., the rate of information transfer) is maximized when food types occur with a particular frequency in the diet. During encounters with novel foods, this "optimal-learning diet" is likely to conflict with classical diet selection theory based on nutrient or energy intake (Belovsky 1981; Stephens and Krebs 1986).

In essence, the process of establishing an association between food types and their postingestive consequences is a process of communication, and the concepts developed in communication theory can provide a useful viewpoint on the problem (Weaver and Shannon 1949; Wells 1999). Conditioning of any kind involves communication, but the application of communication theory to conditioning is still a novel approach (Gallistel 2003). Communication can be divided into three components: emitting a signal, transmitting the signal, and receiving the signal. In the context of food intake and diet selection, each consumed food type emits a set of chemical signals. Each signal is transmitted through a series of metabolic pathways before it eventually elicits a postingestive consequence (the received signal). If an animal is to correctly associate a food type with its consequence, the received signal must be isolated from the background noise of other ingested food types. The degree of isolation is quantified by the mutual information, which measures the information about food types that is transmitted by the postingestive consequences. The information transmitted by any one consequence is

given by the partial mutual information (eq. [2]). A weighted sum over all partial mutual informations gives the total mutual information (eq. [1]). A natural weighting is the relative frequency of each consequence, but other weightings are possible. For example, weightings may reflect the different long-term fitness effects of the postingestive consequences.

An animal can modify the mutual information by changing its feeding behavior (e.g., longer feeding sessions on a single food type) or by changing the rate at which the ingested compounds are metabolized (e.g., changing the nutrient status of an animal). However, for a given process of communication there is a maximum mutual information. This maximum depends on the estimated time delays for each consequence and the frequency of the food types in the diet. This maximum places an upper limit on the rate at which associations can be developed. Furthermore, this maximum is achievable only when food types have a specific relative frequency in the diet, which raises the possibility that during the development of dietary associations, diet selection preferences are modified to increase mutual information. The autocorrelation of the feeding time series may also be modified to match the spectral properties of the transmission process (i.e., the metabolic pathway of the various plant compounds) in order to optimize the performance of the communication system (Wells 1999).

The idea of a communication process is also clear from the viewpoint of a prey species. If a prey species is to defend itself against predation by becoming unpalatable (e.g., by producing secondary metabolites), then its unpalatable message must be evident to a predator. Clear communication between prey and predator is even more important if unpalatability is costly to the prey species. As for the predator, the prey species can also modulate the mutual information. For example, prey spatial distribution can affect a predator's feeding patterns, and more immediate postingestive consequences increase mutual information. Experiments support this time delay effect because shorter delays condition for stronger responses (Mackintosh 1983; Mazur 1994), and with long delays, aversions appear to be more readily elicited than preferences (evidence in rats [Revusky and Garcia 1970; Elizalde and Sclafani 1988; Capaldi and Sheffer 1992] and sheep [Burritt and Provenza 1991; Villalba et al. 1999; Arsenos et al. 2000]).

Communication and signaling are important concepts in other fields of ecology where honesty and mimicry have become important issues (Endler and Basolo 1998; Schaefer et al. 2004). Müllerian mimicry, resulting in plant defense guilds (Atsatt and O'Dowd 1976; Augner 1994), could be advantageous because a common signal has a higher probability of correct association, although recent results show that variation in chemical defenses can in-

crease the rate of learning in chicks and decrease their rate of extinction (Skelhorn and Rowe 2005). Batesian mimicry has also been studied in the context of herbivore defense, with the conclusion that mimicry of plant defense signals could be a common strategy (Augner and Bernays 1998). Batesian mimicry is not necessary for an undefended plant to gain protection because plants susceptible to herbivory can gain protection from spatial association with other plants. Two basic mechanisms have been proposed to explain such protection (Atsatt and O'Dowd 1976): first, a patch with repellent plant species reduces the ability of an herbivore to utilize normal food (repellent-plant hypothesis), and second, an herbivore focuses on a highly preferred plant, reducing the likelihood of ingesting less palatable plants in a patch (attractant-decoy hypothesis). In contrast to the attractant-decoy hypothesis, Wahl and Hay (1995) proposed that less palatable foods may encounter increased herbivory as a result of the presence of palatable species (shared-doom hypothesis). In support of the shared-doom hypothesis, Hjältén et al. (1993) found that herbivory by hares and voles on low-palatability trees was increased if neighboring trees had a high palatability. This may also explain why a toxic plant is more likely to be eliminated from vegetation by generalist herbivory when it is rare than when it is abundant (Provenza et al. 2003). Frequency-dependent foraging has also been reported in woody species, which suffer more severe browsing when they are rare (Lundberg and Danell 1990; Augustine et al. 1998; Edenius et al. 2002). It seems that the outcome of mixing less and more palatable resources does not follow a general rule but may vary with relative abundance and distribution, strength of the signal, and the time delay between ingestion and consequence. Our formulation of the problem in terms of mutual information makes it possible to combine these different factors and give an explanation for the situations in which any one of the above hypotheses is likely to apply.

Foraging behavior arising from incorrect association and the resulting selection pressures on the spatial association between food types may be mirrored by the processes of short-term apparent competition (Holt 1984; Holt and Kotler 1987). By considering a patch containing several food types and a forager whose patch-leaving rule obeys the marginal-value theorem (Charnov 1976), Holt and Kotler (1987) described several methods by which increasing the density of one food type can increase foraging on the other food types. This apparent competition can select for habitat partitioning between food types. One important difference between the apparent competition of Holt and Kotler (1987) and the concept of inaccurate association lies in the timescale of the processes. Apparent competition is strongly influenced by the immediate feeding environment that a forager is experiencing. As a con-

sequence, a forager's food preference can seem to change suddenly as it moves into a new patch. In contrast, food associations are based on past foraging history, which implies that a forager's preference is unlikely to show sudden changes between consecutive foraging patches.

The timescale over which food associations change depends upon the process of learning, although a study of learning is beyond the scope of this article. Our results assume a long memory and give the expected associations based on all the information within the feeding time series. Learning allows the estimates of a consequence's time delay to be updated by new information (e.g., Bayesian updating of the information). An important consideration is the rate at which information is forgotten and the weighting given to new information. Preliminary results using an exponentially weighted scalar memory require the rate of information loss and information gain to be the same. This has implications for the accuracy of an association because if a forager with such a memory is to quickly adapt to new information, it must also quickly forget about existing information. This results in either a forager whose diet selection is insensitive to new information but whose errors of association are small or a forager who adapts quickly to new information but whose errors of association are large because they are based on a small set of recent data.

Our analysis ignores the long-term fitness consequences of an association in order to simplify the analysis by disregarding the sign and strength of a postingestive signal. However, signal strength can be incorporated by extending the model so that a consequence's partial mutual information is weighted according to its contribution to an animal's long-term fitness. This extension is not expected to change our qualitative results, but the incorporation of signal strength into our framework would fuse the ideas on food associations with ideas from classical diet selection theory. The strength of postingestive consequences is known to affect diet preferences (du Toit et al. 1991; Arsenos and Kyriazakis 1999; Arsenos et al. 2000; Villalba and Provenza 2000). For example, lambs and goats can associate the amount of food ingested with the strength of postingestive consequences (Provenza et al. 1994; Villalba and Provenza 2000; Duncan et al. 2006). An animal's physiological state (e.g., nutrient status) is also likely to interact with the strength of a postingestive consequence, modifying its perception of the postingestive consequence.

While the majority of studies on conditioned responses derive from the psychological literature, the consequences of associational learning for ecology and evolution are widely ranging. For example, a stronger association is to be expected between foods and their negative consequences compared with their positive consequences. This is because a food should provide information to the forager about its negative properties to reduce its consumption, whereas from

the food's perspective, its positive properties should remain unclear to the forager to reduce the chances of the food becoming preferred. From the forager's perspective, the degree to which it makes associations between the food and its properties will depend on the impact of the nutrient or toxin on its fitness. Clearly, where the impact is high, the association would be expected to be strong, irrespective of whether the impact is positive or negative. Therefore, the selection pressures are in the same direction for the negative consequences of foods but in opposite directions for the positive consequences of foods. This may be why foragers generally respond more strongly to the negative properties of foods than to the positive ones.

Our analysis demonstrates that various observed diet selection patterns can result from an animal's inability to develop an accurate association between the food types in its diet and their postingestive consequences. The analysis presents testable hypotheses for the relationship between diet associations and food properties. The analysis also complements classical diet selection models. Classical models address the long-term fitness consequences, given perfect knowledge of associations between the feeding environment and its consequences, while this work concentrates on the acquisition of knowledge about the feeding environment. This analysis emphasizes the important role of communication between plant and animal in establishing associations and the applicability of concepts that were initially developed to analyze human telecommunications. Further work is required to test these predictions, to embed food type associations in traditional diet models, and to investigate the extent to which the processes of communication between animals and their dietary food types have coevolved.

Acknowledgments

Literature Cited

Arsenos, G., and I. Kyriazakis. 1999. The continuum between preferences and aversions for flavoured foods in sheep conditioned by administration of casein doses. Animal Science 68:605–616.

Arsenos, G., J. Hills, and I. Kyriazakis. 2000. Conditioned feeding responses of sheep towards flavoured foods associated with casein administration: the role of long delay learning. Animal Science 70:157–169.

Atsatt, P. R., and D. J. O'Dowd. 1976. Plant defense guilds. Science 193:24–29.

Augner, M. 1994. Should a plant always signal its defence against herbivores? Oikos 70:322–332.

Augner, M., and E. A. Bernays. 1998. Plant defence signals and Batesian mimicry. Evolutionary Ecology 12:667–679.

Augustine, D. J., L. E. Frelich, and P. A. Jordan. 1998. Evidence for two stable states in an ungulate grazing system. Ecological Monographs 8:1260–1269.

Bateson, M., and A. Kacelnik. 1998. Risk-sensitive foraging: decision making and variable environments. Pages 2489–2501 *in* R. Dukas, ed. Cognitive ecology: the evolutionary ecology of information processing and decision making. University of Chicago Press, Chicago.

Bautista, L. M., J. Tinbergen, and A. Kacelnik. 1998. Optimal foraging and beyond: how starlings cope with changes in food availability. American Naturalist 152:543–561.

Behmer, S. T., S. J. Simpson, and D. Raubenheimer. 2002. Herbivore foraging in chemically heterogeneous environments: nutrients and secondary metabolites. Ecology 83:2489–2501.

Belovsky, G. E. 1981. Food plant selection by a generalist herbivore: the moose. Ecology 62:1020–1030.

Burritt, E. A., and F. D. Provenza. 1991. Ability of lambs to learn with a delay between food ingestion and consequences given meals containing novel and familiar foods. Applied Animal Behaviour Science 32:179–189.

Capaldi, E. D., and J. D. Sheffer. 1992. Contrast and reinforcement in consumption. Learning and Motivation 23:63.

Carlson, N. R. 1994. Physiology of behavior. 5th ed. Allyn & Bacon, Boston.

Charnov, E. L. 1976. Optimal foraging: attack strategy of a mantid. American Naturalist 110:141–151.

Covasa, M., and J. M. Forbes. 1994. Exogenous cholecystokinin octapeptide in broiler chickens: satiety, conditioned colour aversion and vagal mediation. Physiology and Behaviour 56:39–49.

Crystal, J. D. 2001. Nonlinear time perception. Behavioural Processes 55:35–49.

Darmaillacq, A.-S., L. Dickel, and M.-P. Chichery. 2004. Rapid taste aversion learning in adult cuttlefish, *Sepia officinalis*. Animal Behaviour 68:1291–1298.

Duncan, A. J., and S. A. Young. 2002. Can goats learn about foods through conditioned food aversions and preferences when multiple food options are simultaneously available? Journal of Animal Science 80:2091–2098.

Duncan, A. J., C. Ginane, D. A. Elston, A. Kunaver, and I. J. Gordon. 2006. How do herbivores trade off the positive and negative consequences of diet selection decisions? Animal Behaviour 71:93–99.

du Toit, J. T., F. D. Provenza, and A. Nastis. 1991. Conditioned taste aversions: how sick must a ruminant get before it learns about toxicity in foods? Applied Animal Behaviour Science 30:35–46.

Edenius, L., G. Ericsson, and P. Naslund. 2002. Selectivity by moose vs. spatial distribution of aspen: a natural experiment. Ecography 25:289–294.

Elizalde, G., and A. Sclafani. 1988. Starch-based conditioned flavor preferences in rats: influence of taste, calories and CS-US delay. Appetite 11:179–200.

Endler, J. A., and A. L. Basolo. 1998. Sensory ecology, receiver biases and sexual selection. Trends in Ecology & Evolution 13:415–420.

Farnsworth, K. D., and A. W. Illius. 1998. Optimal diet choice for large herbivores: an extended contingency model. Functional Ecology 12:74–81.

Forbes, J. M. 1999. Natural feeding behaviour and feed selection. Pages

3–12 *in* D. van der Heide, E. A. Huisman, E. Kanis, J. W. M. Osse, and M. W. A. Verstegen, eds. Regulation of feed intake. CABI, Oxford.

Fraser, A. M., and H. L. Swinney. 1986. Independent coordinates for strange attractors from mutual information. Physical Review A 33: 1134–1140.

Gallistel, C. R. 2003. Conditioning from an information processing perspective. Behavioural Processes 62:89–101.

Garcia, J., and W. G. Hankins. 1977. On the origin of food aversion paradigms. Pages 1–19 *in* L. M. Barker, M. Best, and M. Domjan, eds. Learning mechanisms in food selection. Baylor University Press, Waco, TX.

Garcia, J., and D. J. Kimeldorf. 1960. Conditioned avoidance behavior induced by low-dose fast neutron exposure. Nature 185:261.

Garcia, J., F. R. Ervin, and R. A. Koelling. 1966. Learning with prolonged delay of reinforcement. Psychonomic Science 5:121–122.

Gelperin, A. 1975. Rapid food-aversion learning by a terrestrial mollusk. Science 189:567–570.

Grimmett, G. R., and D. R. Stirzaker. 1992. Probability and random processes. 2nd ed. Clarendon, Oxford.

Gustavson, C. R., D. J. Kelly, M. Sweeney, and J. Garcia. 1976. Prey lithium aversions. I. Coyotes and wolves. Behavioural Ecology 17: 61–72.

Hjältén, J., K. Danell, and P. Lundberg. 1993. Herbivore avoidance by association: vole and hare utilization of woody plants. Oikos 68:125–131.

Holt, R. D. 1984. Spatial heterogeneity, indirect interactions, and the coexistence of prey species. American Naturalist 124:377–406.

Holt, R. D., and B. P. Kotler. 1987. Short-term apparent competition. American Naturalist 130:412–430.

Houston, A. I. 1993. The importance of state. Pages 10–31 *in* R. N. Hughes, ed. Diet selection: an interdisciplinary approach to foraging behaviour. Blackwell, Oxford.

Houston, A. I., and J. M. McNamara. 1999. Models of adaptive behaviour: an approach based on state. Cambridge University Press, Cambridge.

Kyriazakis, I., B. J. Tolkamp, and G. Emmans. 1999. Diet selection and animal state: an integrative framework. Proceedings of the Nutrition Society 58:765–772.

Lundberg, P., and K. Danell. 1990. Functional response of browsers: tree exploitation by moose. Oikos 58:378–384.

Mackintosh, N. J. 1983. Conditioning and associative learning. Oxford Psychology Series. No. 3. Oxford University Press, New York.

Mazur, J. E. 1994. Learning and behavior. Prentice Hall, Englewood Cliffs, NJ.

Newman, J. A., A. J. Parsons, J. H. M. Thornley, P. D. Penning, and J. R. Krebs. 1995. Optimal diet selection by a generalist grazing herbivore. Functional Ecology 9:255–268.

Paninski, L. 2003. Estimation of entropy and mutual information. Neural Computation 15:1191–1253.

Pavlov, I. P. 1927. Conditioned reflexes. Oxford University Press, Oxford.

Provenza, F. D. 1995. Postingestive feedback as an elementary determinant of food preference and intake in ruminants. Journal of Rangeland Management 48:2–17.

Provenza, F. D., J. J. Lynch, E. A. Burritt, and C. B. Scott. 1994. How goats learn to distinguish between novel foods that differ in post-ingestive consequences. Journal of Chemical Ecology 20:609–624.

Provenza, F. D., J. J. Villalba, and J. P. Bryant. 2003. Foraging by herbivores: linking the biochemical diversity of plants with herbivore culture and landscape diversity. Pages 387–421 *in* J. A. Bissonette and I. Storch, eds. Landscape ecology and resource management: linking theory with practice. Island, New York.

Pulliam, H. R. 1974. On the theory of optimal diets. American Naturalist 108:59–74.

Revusky, A., and J. Garcia. 1970. Learning associations over long delays. Pages 1–84 *in* G. H. Bower, ed. The psychology of learning and motivation. Academic Press, New York.

Schaefer, H. M., V. Schaefer, and D. J. Levey. 2004. How plant-animal interactions signal new insights in communication. Trends in Ecology & Evolution 19:577–584.

Schoener, T. W. 1971. Theory of feeding strategies. Annual Review of Ecology and Systematics 2:369–404.

Sclafani, A. 1995. How food preferences are learned: laboratory animal models. Proceedings of the Nutrition Society 54:419–427.

Skelhorn, J., and C. Rowe. 2005. Tasting the difference: do multiple defence chemicals interact in Müllerian mimicry? Proceedings of the Royal Society of London B 272:339–345.

Stephens, D. W. 1989. Variance and the value of information. American Naturalist 134:128–140.

Stephens, D. W., and J. R. Krebs. 1986. Foraging theory. Princeton University Press, Princeton, NJ.

Susswein, A. J., M. Schwarz, and E. Feldman. 1986. Learning changes of feeding behaviour in *Aplysia* in response to edible and inedible foods. Journal of Neuroscience 6:1513–1527.

van Marken Lichtenbelt, W. D. 1993. Optimal foraging of a herbivorous lizard, the green iguana in a seasonal environment. Oecologia (Berlin) 95:246–256.

van Wieren, S. E. 1996. Do large herbivores select a diet that maximizes short-term energy intake rate? Forest Ecology and Management 88:149–156.

Villalba, J. J., and F. D. Provenza. 1999. Nutrient-specific preferences by lambs conditioned with intraruminal infusions of starch, casein and water. Journal of Animal Science 77:378–387.

———. 2000. Roles of flavour and reward intensities in acquisition and generalization of food preferences: do strong plant signals always deter herbivory? Journal of Chemical Ecology 26:1911–1922.

Villalba, J. J., F. D. Provenza, and J. Rogosic. 1999. Preference for flavoured wheat straw by lambs conditioned with intraruminal infusions of starch administered at different times after straw ingestion. Journal of Animal Science 77:3185–3190.

Wahl, M., and M. E. Hay. 1995. Associational resistance and shared doom effects of epibiosis on herbivory. Oecologia (Berlin) 102: 329–340.

Weaver, W., and C. E. Shannon. 1949. The mathematical theory of communication. University of Illinois Press, Urbana.

Wells, R. B. 1999. Applied coding and information theory for engineers. Prentice Hall, Upper Saddle River, NJ.

Westoby, M. 1974. An analysis of diet selection by large generalist herbivores. American Naturalist 108:290–304.