Risky Choice and Weber’s Law

ALEX KACELNIK* AND FAUSTO BRITO E ABREU

Department of Zoology, University of Oxford, Oxford OX1 3PS, U.K.

(Received on 13 February 1998, Accepted in revised form on 8 June 1998)

We present a family of models of choice between behavioural alternatives with stochastic outcomes (risky choice) based on the effects of Weber’s Law in memory. These models generalise and extend a model of risk sensitive foraging originally proposed by Reboreda & Kacelnik [(1991) Behav. Ecol. 2, 301–308], which yielded qualitative predictions (risk-aversion for amount of food and risk-proneness for delay to food). We now demonstrate how this approach can predict quantitatively the partial preferences between two alternative options with any mean and variance in their outcomes, and the certainty equivalent of an option consisting of any set of probabilistic outcomes. The approach is also relevant to the economics and psychology of risk sensitivity because it predicts risk aversion for any desirable outcome (such as monetary gains) and risk seeking for any undesirable gain (such as monetary losses). Our models are process-based rather than purely normative, and are based on linear expected utility as a function of expected outcomes. They do not account for all observed aspects of risky choice, but their descriptive performance betters that of existing functional models and requires fewer parameters.

1. Introduction

We generalise and extend a model of choice between behavioural alternatives with stochastic outcomes (risky choice). The ideas leading to this contribution originate most directly on previous work by Gibbon et al. (1988), Reboreda & Kacelnik (1991), Bateson (1993), Bateson & Kacelnik (1995b) and Brunner et al. (1994, 1996). Here we formalise predictions of the model for a broader class of choices, give details of how the model can predict quantitatively the proportion of choices between alternative options and separate the essential from ancillary assumptions of the model by showing that the original model is a member of a broader class of models that share a common main assumption: Weber’s Law. We start by introducing the problem of risky choice.

When animals (including humans) are faced with behavioural options that deliver the same average payoffs but differ in variance, expected reward (we use “expected” to refer to the arithmetic mean, or first moment of the distribution of outcomes) does not account fully for observed choice: differences in the variance associated with each option play a major role on preferences (for recent reviews of experimental evidence see Kacelnik & Bateson, 1996, 1997).

In most studies of risky choice using non-human subjects, experimenters have looked at preference between alternatives that differ in either variability in size of food rewards or in variability in delay between choice and outcome.
Almost universally, the results show that the sign of preference can be explained by the dimension responsible for the variance. When variance is introduced controlling the food amounts delivered by each alternative (e.g. food reward size or number of reward items) animals are most often risk averse and rarely risk prone. Conversely, when variance depends on the delays to food delivery (e.g. search or handling time) animals are risk prone.

Many theoretical treatments of risky choice in non-humans have addressed the fitness consequences of variance in payoff (see review by McNamara & Houston, 1992 and several contributions to the recent special issue on risk sensitivity in American Zoologist Vol. 36, 1996), identifying for each theoretical scenario whether higher variance or lower variance in outcomes would result in higher expected fitness. These models face several empirical challenges, of which the two most salient are the following. First, they do not predict the weak but systematic bias towards risk aversion for amounts and universal risk proneness for delays. Some previous work comes close to this: McNamara & Houston (1987) demonstrate that delays and amounts pose different problems for risk sensitivity theory and McNamara (1996) shows that environmental conditions favouring risk aversion for amount are likely to be more frequent than those favouring risk seeking, while various functional models show that risk proneness for delays could be a consequence of time discounting being nonlinear and rate computation being based on local rather than global estimates (Mazur, 1984; Bateson & Kacelnik, 1996). Second, functional models per se, without any provision for errors, do not predict partial preferences: the option yielding higher expected fitness ought to be preferred exclusively. Since amount and delay variance have such conspicuous and systematic different effects, and partial preferences are ubiquitous in choice, these theories need special extensions (not considered here) to accommodate the data. An additional empirical difficulty with these functional models of risky choice is that their predictions are dependent on various parameters (such as those determining the shape of the fitness versus payoff function) which are inaccessible to research, most notably the time base over which the fitness consequences of series of choices should be integrated.

Reboreda & Kacelnik (1991) followed a process-based approach, using developments in the theory of time perception (Gibbon, 1977; Gibbon et al., 1988, 1984) which they extended to the perception of food amounts. In the next section we explain, expand and discuss their model and later we show that their central assumptions can generate a broader family of models with the same major properties.

The rationale core of our approach can be summarised in the following points:

- the empirical finding in the non-human data that must be accommodated by theoretical models of risky choice is that when alternatives have equal mean payoff, subjects favour alternatives with lower variance in size of payoff and higher variance in delay to reward. In the human literature, the data show trends to risk aversion for gains and risk proneness for losses (Kahneman & Tversky, 1979). Hence, the process of choice must be controlled by some statistic of outcome distributions that fosters risk avoidance for desirable outcomes and risk seeking for aversive outcomes;
- the statistic of outcome distributions is to be found in the path between real outcomes of past behaviour and choice between future actions, namely in the process by which subjects perceive, represent, retrieve and use their experience to generate behaviour;
- modelling this process must be restricted to hypotheses that accommodate present knowledge about perception and memory;
- since neither the mean nor (symmetrical) variance of payoff distributions introduce a bias against variability in gains and in favour of variability in losses as required in our first point, we search for the culprit in the third moment (skew).

In the following section we describe our implementation of these points as a theoretical core for the approach, and later on we differentiate specific models.
2. From Symmetric Payoffs to Skewed Representations: Weber’s Law

Our crucial step is to apply Weber’s Law to the subjective representation of time intervals and of amounts of food. Weber’s Law (or Weber–Fechner Law) is a nearly universal empirical finding in psychophysics, and has been described in a variety of forms. In discrimination tests, the law states that the just noticeable difference (JND) between a test stimulus and a standard is proportional to the value of the latter. Thus, if given a standard weight of 10 g a subject cannot detect that a comparison is lighter or heavier unless it is below 9 g or above 11 g, respectively, for a standard weight of 10 kg the required comparisons will be 9 or 11 kg. In production tests, when the subject repeatedly generates an output of a desired (or instructed) magnitude, Weber’s Law implies that the displayed behaviour is normally distributed, with standard deviation proportional to the mean magnitude of the response. This is particularly clear in behavioural timing. If a subject is repeatedly rewarded $x$ seconds after a signal, and occasionally induced to show the time at which the reward is anticipated (as in the peak procedure, Catania, 1970; Roberts, 1981) the distribution of the time of responding is normally distributed around $x$, with constant coefficient of variation as $x$ is modified between conditions. It is a matter of debate whether this observation, which in behavioural timing is known as the scalar property (Gibbon, 1977, 1991), is generated by error in perception, in internal representation (memory), or in producing behaviour.

In the present implementation, the subject’s knowledge of behavioural outcomes (the food amounts or time intervals that result from each action) is modelled as a set of normal distributions with mean and standard deviation proportional to each experienced outcome. When a behavioural option yields more than one possible outcome, each with a given probability, the subject’s representation of that option is the sum of the normal distributions generated by each of the outcomes, with areas proportional to their respective probabilities (see Fig. 1). The coefficient of variation of these distributions (the proportionality constant between the mean and the standard deviation) is assumed to be independent of payoff magnitude. The assumption of constant coefficient of variation (a strong form of Weber’s Law) is founded on an extensive empirical literature on memory for time intervals (for a review see Gibbon, 1991) and on more limited evidence for amounts of food (Bateson & Kacelnik, 1995a).

Because of Weber’s Law, larger outcomes are represented internally less sharply than smaller ones (in absolute terms). As a result, the sum of the representation of the outcomes of a variable option is distorted with respect to the veridical distribution of outcomes. If, for example, a behaviour causes a uniform distribution of outcomes (several possible outcomes all with equal probability), the corresponding subjective representation is not uniform but positively skewed, because larger outcomes produce flatter distributions than smaller ones (Gibbon et al., 1988). In general, memory distributions will always be more positively skewed than payoff distributions (in Fig. 1 we show a case where a negatively skewed distribution of outcomes generates a positive skewed distribution of its representation). Between uniform distributions of outcomes, the greater the variance the stronger the positive skew of the corresponding internal representation. The logic of the model is described in Fig. 1.

Notice that there is a fundamental difference in how utility is treated in our approach and that of either risk sensitivity theory in biology, prospect theory in psychology and utility theory in microeconomic handling of risk aversion. The present model assumes a linear representation of payoff and a monotonic assignment of value. The alternative approaches (for an exception see Lopes, 1995) are fully dependent on the curvature of the value versus payoff function.

The assumptions described so far we believe to be strongly based on empirical evidence, and we shall keep it as a common base for several possible specific implementations. The main result is that because of a fundamental property of cognitive architecture, variability induces a shift towards more positive skews in the subject’s knowledge of its opportunities. The specific models we discuss in the following sections use this feature to generate the observed biases in preference.
Fig. 1. Schematic representation of the proposed mechanism for choice between a fixed and a variable option. In this example the fixed option has an outcome of five units with certainty (\(P = 1\)), while the variable option has the same mean pay-off but three possible outcomes: two units (\(P = 0.3\)), four units (\(P = 0.3\)), and eight units (\(P = 0.4\)). The six distributions shown have magnitude on the X axis and probability on the Y axis, the latter ranging from the probability of real outcomes (top), through that of the way these outcomes are perceived or stored as affected by Weber’s Law (middle) to the aggregate representation of the expected outcome of the behaviour that elicits the respective outcomes (bottom). In the equations, \(\Phi\) is the cumulative distribution function of the internal representation’s probability density functions (\(f_F\) and \(f_V\)).
3. Single Sampling and Partial Preferences

3.1. TWO EQUIPROBABLE OUTCOMES

In addition to modelling the subject’s knowledge, it is necessary to postulate a decision process. Gibbon et al. (1988) and then Reboreda & Kacelnik (1991) implemented choice as a competition between samples from the representations of expected outcomes. To generate each action the decision-maker picks a sample from its internal representation of the outcomes of each behaviour and performs the behaviour associated with the most favourable sample (i.e. the larger amount or the shorter time interval).

The most common case considered in the empirical literature on risky choice refers to preferences between a “fixed” option that provides a single outcome with certainty and an alternative (“variable”) option that has a set of two equiprobable outcomes with mean equal to the value of the fixed option. In this case, because of the symmetry of the representation of the fixed option and the positive skew of the variable one, the latter yields a smaller sample in the majority of comparisons—a shorter delay or a smaller amount. Therefore, in agreement with the empirical evidence, this model predicts a majority of choices of the variable option (risk-seeking) when variance is due to delays (the smaller sample is preferred because delays are aversive) and a minority of choices of the variable option (risk aversion) when variance is due to amounts of reward (the bigger sample is preferred because larger food amounts are desirable).

Reboreda & Kacelnik (1991) used this model to produce this qualitative prediction, but later Bateson (1993) and Bateson & Kacelnik (1995b) computed the certainty equivalent of a variable alternative with the same expected value. She showed that in such scenarios preference is a bitonic function of outcome variance: when the two outcomes of the variable option are either so far apart that their representations show little overlap or so close that they produce little skew the subject approaches indifference, while intermediate values cause maximum bias. The empirical evidence for this prediction is still inconclusive.

3.2. ARBITRARY SETS OF OUTCOMES

In this section, we calculate this model’s predictions for the certainty equivalent and for partial preferences in choices between fixed options and variable options with arbitrary sets of outcomes. Moreover, we show that this model can also be used to predict the proportion of choices between two variable options (or two fixed options), regardless of their mean and variance. These generalisations make more exhaustive quantitative tests of the model possible, and should help in probing the generality of the account it offers for different scenarios of non-human and human cases of risky choice.

According to the model, when a subject learns about an alternative yielding a fixed outcome, its memory representation is a function of the outcome magnitude \(m_F\) and of a subject-specific parameter \(\gamma\), which represents the coefficient of variation in the internal representation of outcomes. We represent a sample \(S_F\) taken from the probability density function \(f_F\) in the animal’s internal representation by

\[
S_F \sim f_F = N(\mu_F, \gamma \cdot \mu_F) \tag{1}
\]

where \(N(\mu, \sigma)\) is a normal distribution with mean \(\mu\) and standard deviation \(\sigma (\sigma = \gamma \cdot \mu)\). Similarly,
the probability density function describing the representation of a variable option is a function of the different outcomes experienced in that option \((\mu_1, \mu_2, \ldots, \mu_n)\), their respective probabilities \((P(V_1), P(V_2), \ldots, P(V_n))\), and the animal’s \(\gamma\). We can represent a sample \((S_V)\) taken from the reference memory for the variable option by

\[
S_V \sim f_V = \sum_{i=1}^{n} P(V_i) \cdot N(\mu_v; \gamma, \mu_v
\]

where \(f_V\) is the probability density function describing the internal representation of the variable option, and \(n\) is the number of different outcomes of that option.

In cases when the hedonic value of different outcomes is a positive function of magnitude, such as when dealing with amounts of food, the subject chooses the fixed option whenever \(S_V < S_F\), and the variable option when \(S_V > S_F\). In such cases, the probability of the animal choosing the variable option equals the probability of picking a given value of \(S_V\) (given by the variable option’s probability density at \(S_V\)) and simultaneously picking a smaller value of \(S_F\). The chances that a random sample taken from \(f_F\) is smaller than any particular number is given by the cumulative distribution of \(f_F\) up to that number. Since the numbers of interest are all possible values of \(S_V\) the chances of \(S_V > S_F\) are given by the integral shown in eqn (3).

\[
P(\text{choose Var.}) = \int_{0}^{+\infty} f_V(x) \cdot P(S_V < x) \, dx
\]

where \(x\) is the integration variable (the possible values taken by \(S_V\)). Notice that although the model assumes normal distributions, the integral is evaluated only for positive values of \(x\) because negative values for food amounts and for time delays are unrealistic and can be ignored. Equation (3) can be rewritten as

\[
P(\text{choose Var.}) = \int_{0}^{+\infty} f_V(x) \cdot \Phi_V(x) \, dx
\]

where \(\Phi_V\) is the cumulative distribution function of \(f_V\).

We can combine the expressions in (1), (2) and (4) to compute the probability of choosing the variable option—which in turn allows us to predict the percentage of choices for each option in a given experiment—using the following variables: the magnitude of the outcomes in the two options \((\mu_F, \mu_{V1}, \mu_{V2}, \ldots, \mu_{VN})\), the probability of each of the outcomes of the variable option \((P(V_1), P(V_2), \ldots, P(V_N))\), and the individual’s coefficient of variation \((\gamma)\).

Notice that since outcome magnitudes and probabilities can be known with certainty because they are programmed by the experimenter, and \(\gamma\) can be estimated by a variety of experimental techniques, the model has no free parameters. To find the certainty equivalent of a variable option (the fixed outcome \(\mu_F\) of the fixed alternative that would be chosen equally often) we set \(P(\text{choose Var.}) = 0.5\) and solve for \(\mu_F\) using a symbolic solution software package such as Mathematica (Wolfram, 1992).

All these calculations apply equally well to outcomes whose variance is manipulated as food amounts or as delays, the only difference being that in (3) the probability of choosing variable when value is a decreasing function of outcome magnitude (as when delay to food is manipulated) is \(P(S_V < S_F)\), while for desirable outcomes such as food amounts it is \(P(S_V > S_F)\).

It should be noted that eqn (1) is a particular case of eqn (2) [with \(n = 1\) and \(P(V) = 1\)]. With the appropriate adjustments eqn (4) can also be used to predict the proportion of choices for two variable options (or two fixed options) with any means and variances.

Table 1 shows a comparison of the predictions of eqn (4) and the results of studies by Reboreda & Kacelnik (1991) and Bateson & Kacelnik (1995b) on risk sensitivity of starlings. Table 1(a) refers to the results of the treatments where variance was introduced in the amount of food. In both studies the model predicts a small degree of risk aversion, which is also shown in the data. The results do not differ significantly from the model’s predictions. Table 1(b) shows the results of the treatment where there was variability in time delays. In this case eqn (4) predicts a small degree of risk proneness, but the results, while having the appropriate sign, are more extreme, and differ significantly from the predictions.
Details of the studies by Reboreda & Kacelnik (1991) and Bateson & Kacelnik (1995b) which tested starlings’ preference for risk in both food amount (a) and delay to reward (b). The table describes the number of birds used in each study (n), the outcomes delivered by the fixed and variable options, the observed percentage of choices for the variable option, and the predicted proportion of choices for the variable option computed with eqn (4). Column “p” shows the p value of a Wilcoxon’s signed-rank test of the results of each study against the hypothesised mean predicted by the model.

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Fixed option</th>
<th>Variable option</th>
<th>% choices variable</th>
<th>Prediction eqn (4)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reboreda &amp; Kacelnik (1991)</td>
<td>12</td>
<td>4</td>
<td>2.7 or 5.3</td>
<td>42.9</td>
<td>47.6</td>
<td>0.606</td>
</tr>
<tr>
<td>Bateson &amp; Kacelnik (1995b)</td>
<td>6</td>
<td>5</td>
<td>3 or 7</td>
<td>48.2</td>
<td>46.6</td>
<td>0.688</td>
</tr>
</tbody>
</table>

We have selected these particular studies for two reasons. First, they compare the effects of variance in amount and delay, that is a desirable and an aversive dimension, in the same experimental set up. Second, there is published information on γ in starlings for both amounts of food and time intervals. There is considerable individual variation in this species, and it ranges from 0.16 to 0.6 for time (Brunner et al., 1992), and from 0.3 to 0.8 for amount (Bateson & Kacelnik, 1995a). Our predictions were made using γ = 0.6 for amounts and γ = 0.4 for delays. However, the predictions are relatively insensitive to variations in γ within those ranges and a sensitivity analysis showed that exactly the same picture is produced by using the extreme values of individual variation.

The main quantitative failure of the model observed so far is its inability to predict the extreme risk proneness for delay found in empirical studies. We believe that this could be caused by synergy with other factors that independently favour risk proneness for delays. One (mechanistic) example is the conditioning process by which animal subjects are trained to choose: the strength of the association between a manipulandum such as a pecking key and a delayed reward is likely to be an accelerated function of temporal contiguity. Another (functional) factor is the set of various reasons leading to nonlinear time discounting combined with local computation of average rates. These ideas are discussed elsewhere (Bateson & Kacelnik, 1996; Kacelnik, 1997; Kacelnik & Bateson, 1996, 1997).

The fit of the model to data could be improved by adding a free parameter to amplify the effect of probabilities as they depart from 50%. This is a usual tool in model development (see for example the generalised matching law in experimental psychology, Baum, 1974; Davison & McCarthy, 1988), but in this case we prefer to avoid it because improving the fit using free parameters may actually obscure discrepancies that may have important heuristic value. The model successfully predicts risk aversion for desirable outcomes such as food amounts or monetary gains and risk seeking for aversive outcomes such as food delays or monetary losses, but while the quantitative predictions for amounts are acceptable, those for delay are significantly less extreme than those shown by real subjects.

4. Alternatives to Single Sampling from Memory

An important weakness of this version of the model is that it relies on the subject retrieving single samples from its internal representation for each choice. This would be surprising, for both functional and empirical reasons. Functionally, this form of sampling and decision leads to
submaximal rate of reinforcement because an option may be preferred even when its mean payoff is lower than the alternative. Thus, it is hard to see why it should be employed if the subject has access to the central values of the distributions. Little is known about the availability of the mean of the representation of variable sets of stimuli to choosing subjects. However, in so-called reproduction tasks, in which subjects display a behaviour that reflects the internal representation of experienced delays, subjects produce within single trials functions with the properties of the functions depicted as \( f_i \) and \( f_r \) in Fig. 1, rather than producing a single value per trial (Bateson & Kacelnik, 1997; Brunner et al., 1996, 1994). Why, having the capacity to map the whole distribution in each trial, should they choose according to single sampling is not at all clear. An additional problem is that while the model predicts that the certainty equivalent of mixtures of two equiprobable outcomes should be at the geometric mean, at least for delays it is well established that certainty equivalents in that case are close to the harmonic mean. This model can accommodate this result by making the additional assumption that the representation of the fixed option has no variance, rather than following Weber’s Law (Gibbon et al., 1988), but there do not seem to be independent reasons to defend this view.

However, the problematic assumption of single sampling from memory is not the only possibility. Any model which used a skew-sensitive statistic of the internal representation would cope with the main features of the data.

For instance, a subject that made its choices using the median (or mode) of the internal representation, will show the appropriate biases for amount and delay. Inspection of Fig. 1 shows this. If the subject chose by comparing the mode of \( f_i \) versus that of \( f_r \), in the example the former would yield a smaller value than the latter. If the dimension of the stimuli were time, the variable set would be preferred, if it were amounts, the opposite would happen. Using the median would yield a similar result.

In contrast with the single sample approach, any approach based on a single statistic of the central value of the distributions would not generate partial preferences. Choice proportions would be all or none. This all-or-none behaviour is a common curse of functional models, and a frequent solution is to suggest that the generation of behaviour from internal representation of outcomes is itself a stochastic process, with probabilities that match the relative payoffs of the alternatives. In operant psychology, the Matching Law (Davison & McCarthy, 1988) implies that the proportion of choices is determined by the ratio between experienced rates of reinforcement, sometimes weighted by free parameters to accommodate experimental results. In the foraging literature a similar example (called profitability matching) is the prediction of the proportion of visit lengths to depleting patches on the bases of the relative profitability that would accrue from each visit length (Kacelnik, 1984). The notion of matching is compatible with the core of our model, because we can postulate a parameter-free choice process according to eqn (5).

\[
P \text{(choose Var.)} = \frac{\text{Median} (f_i)}{\text{Median} (f_r) + \text{Median} (f_i)}
\]

In Fig. 2 we show the predicted partial preferences of single sampling [eqn (4)], and Median matching [eqn (5)], and Mode matching...
[given by eqn (5) replacing medians by modes] for choices between a variable option with the distribution of outcomes presented in Fig. 1 versus a range of fixed options (including the one shown in Fig. 1). The Median and Mode matching models predict risk proneness for delays and risk aversion for amounts, as the original model (with single sampling) does: when the fixed option offers the same expected payoff as the variable option (five units) these two models predict that fixed amounts or variable delays should be preferred (Fig. 2). Moreover, according to these models the certainty equivalent for the variable option (i.e. the value of the fixed option that is chosen the same number of times as the variable option) should be smaller than its expected pay-off (five units).

Similarly to the single sampling model, the Median and Mode matching models are based on the well established effects of Weber’s Law in memory and do not have free parameters. Their predictions are particularly insensitive to $\gamma$, which affects the shape of the internal representations of the two options ($f_v$ and $f_\delta$).

Median and Mode matching are just two examples of models that consider a decision process different from comparing single samples taken at random from the internal representations of the two options. In principle, any model of risky choice based on Weber’s Law, can predict risk proneness for delays and risk aversion for amounts, provided we assume a decision process that is sensitive to the skew of the internal representations of the behavioural alternatives.

5. Conclusion

We have shown how a family of process-based models can account for the general trend of empirical results of studies of risky choice in humans and non-humans. In these studies, subjects tend to be weakly risk averse when there is variance in amount of positive rewards, and strongly risk prone when there is variance in negative outcomes. The models we presented are based on the principle that perceptual or processing errors, when representing dimensions of reward such as amounts of food or time intervals, are proportional to the magnitude of those stimuli—Weber’s Law. They do not include any nonlinearity in utility as most psychological models of economic choice do (see for instance Kahneman & Tversky, 1979; in contrast with the linear utility, process-based model of Lopes, 1995).

In their present form these models have only one (empirical) parameter, namely the coefficient of variation ($\gamma$) in the internal representation of experienced outcomes. While this parameter is not directly accessible to measurement, its value can be estimated using a range of experimental procedures, and it should not be seen as a free parameter to be adjusted to data for better fit. Further, most predictions of the models are rather insensitive to the parameter value.

The approach we present in its present form cannot account for all aspects of risky choice, notably the state dependent switch in preference predicted by functional models of risk sensitivity (for the foundations of the theoretical bases of these predictions see Caraco et al., 1980; McNamara & Houston, 1992; and for a critical review of theory and evidence in support of this prediction see Kacelnik & Bateson, 1996). The descriptive performance of our model, however, betters that of existing functional models in handling the contrast in response to variability in positive (amount) and negative (delay) dimensions of payoff. Further, and its process-based structure makes it heuristically more powerful than competing descriptive accounts because the assumptions can be tested independently.

This research was funded by the Wellcome Trust, U.K. (Grant 046101 to AK). FBA was supported by the Junta Nacional de Investigação Científica e Tecnológica, Portugal (Grant PRAXIS XXI/BD/5870/95) and The Queen’s College, Oxford. We are grateful to Juan Reboreda, John Gibbon, Dani Brunner and Melissa Bateson, who participated in the genesis and development of these ideas, and to Renée Menezes and Ruy Ribeiro for helpful discussions.

REFERENCES


