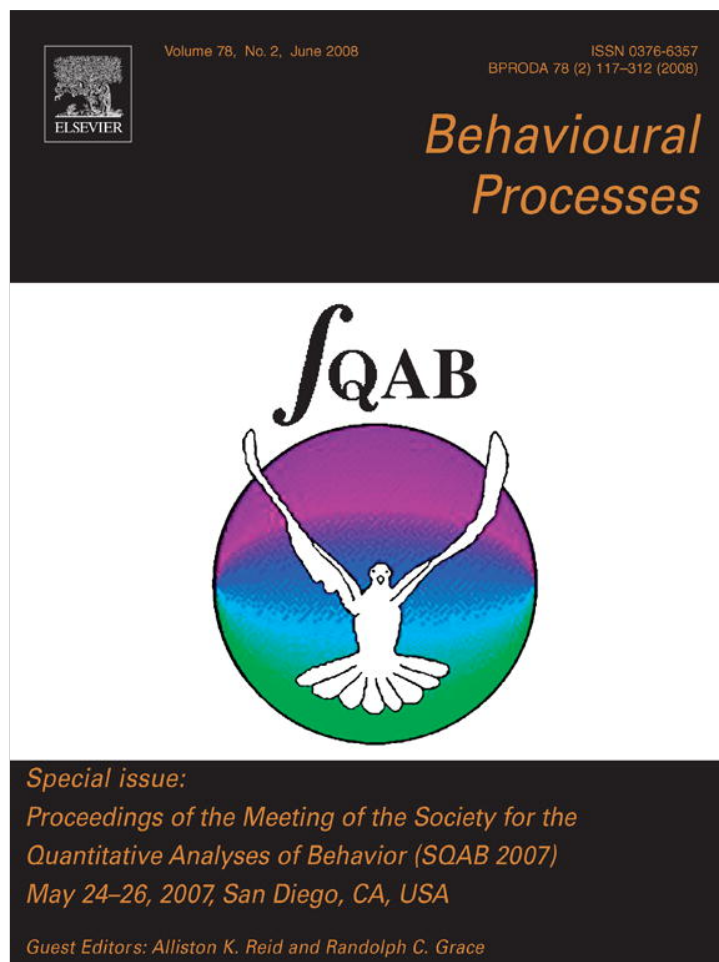


Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



## Short communication

## Matching by fixing and sampling: A local model based on internality

Vladimir A. Lefebvre<sup>a,\*</sup>, Federico Sanabria<sup>b</sup><sup>a</sup> School of Social Sciences, University of California, Irvine, CA 92697, United States<sup>b</sup> Arizona State University, United States

Received 9 September 2007; accepted 4 January 2008

**Abstract**

Undermatching and overmatching in concurrent schedules of reinforcement have been traditionally described as changes in the slope of the Generalized Matching Law function. More recently, it has been suggested that deviations from strict matching may be better described as following a policy of mostly *fixing* on the preferred schedule, and occasionally *sampling* the alternative schedule. So far, no model of local performance predicts the global outcome of this policy. We describe one such model; it assumes immediate and long-term effects of reinforcement on local performance. The model assumes long-term effects as changes in the internal state of the organism. Formally, the model is analogous to the Axiom of Repeated Choice [Lefebvre, V.A., 2004. Bipolarity, choice, and entro-field. In: Proceedings of the Eighth World Multi-Conference on Systemics, Cybernetics and Informatics, vol. IV, pp. 95–99].

© 2008 Elsevier B.V. All rights reserved.

**Keywords:** Concurrent schedules; Choice; Generalized Matching Law; Internality; Visit patterns; Fix-and-sample; Axiom of Repeated Choice; Contingency-discriminability model

The Matching Law is a fundamental concept in behavioral choice theory (Herrnstein, 1961; for a review, see Davison and McCarthy, 1988). It states that the allocation of behavior across alternatives matches the distribution of obtained reinforcers (De Villiers and Herrnstein, 1976); mathematically, it may be expressed as

$$\frac{B_1}{B_1 + B_2} = \frac{r_1}{r_1 + r_2}, \quad (1)$$

or in its equivalent ratio form (Baum and Rachlin, 1969):

$$\frac{B_1}{B_2} = \frac{r_1}{r_2}, \quad (1')$$

where  $B_1$  and  $B_2$  are the rates of responding on each alternative, and  $r_1$  and  $r_2$  are the rates of reinforcement obtained from the corresponding alternatives.

Extensive empirical research on animal and human choice has shown that experimental subjects systematically deviate from “strict” matching (Davison and McCarthy, 1988). To account for these deviations, Baum (1974) suggested a generalization of

Eq. (1') known as the Generalized Matching Law (GML):

$$\log \frac{B_1}{B_2} = a \log \frac{r_1}{r_2} + \log b. \quad (2)$$

Parameter  $a$  represents the *sensitivity* of choice to relative rate of reinforcement (Lobb and Davison, 1975), and  $b$  is the *bias* towards schedule 1 that is not due to  $r_1$  or  $r_2$ . If  $a > 1$ , rates of reinforcement are *overmatched* by choice; if  $a < 1$ , rates of reinforcement are *undermatched*. Undermatching is prevalent across studies (Davison and McCarthy, 1988), although overmatching is occasionally observed (Aparicio, 2001).

Much work has been done in the last three decades to specify local processes that would yield matching in the long run (Davison and Jenkins, 1985; Herrnstein, 1982; MacDonall, 1999; Staddon and Motheral, 1978; Wearden, 1983). Consistent with the global pattern described by GML (Eq. (2)), these local models assume that the function that relates choice to reinforcement – the *matching function* – is continuous. Recent evidence, however, has questioned this assumption.

### 1. Discontinuity in the matching function: fix-and-sample patterns of choice

The Matching Law has been typically demonstrated using concurrent variable-interval variable-interval (Conc VI VI)

\* Corresponding author. Tel.: +1 949 351 8436.  
E-mail address: valefebv@uci.edu (V.A. Lefebvre).

schedules of reinforcement. In this procedure, subjects continuously choose between two alternatives, each of which delivers rewards at a programmed rate but with no periodicity. By examining local patterns of choice in Conc VI VI, Baum et al. (1999) uncovered a “fix-and-sample” pattern in pigeon’s choices: Pigeons generally responded on the alternative that yielded more reinforcers (*rich* schedule) and occasionally made a few responses on the alternative that yielded fewer reinforcers (*lean* schedule); whereas responding on the rich schedule was sensitive to relative rate of reinforcement, responding on the lean schedule was not. This fix-and-sample pattern has also been detected in rhesus monkeys (Lau and Glimcher, 2005) and rats (Aparicio and Baum, 2005).

The fix-and-sample pattern implies a discontinuity in the matching function where schedules change from lean to rich and vice versa. To illustrate this point, consider what happens when a schedule switches from being lean to being rich. While the schedule was lean, the duration of each visit was fixed, say, to four keypecks on the average, regardless of its relative rate of reinforcement. Now that it is the rich schedule, the duration of each visit is sensitive to rate of reinforcement. When the rich schedule is only slightly richer than the lean schedule, sensitivity to rate of reinforcement may drop visit lengths below four keypecks or may boost them above that number, yielding respectively a step down or up in the matching function.

To account for the fix-and-sample pattern observed in their experiment and for the discontinuity that it implies for the matching function, Baum et al. (1999) suggested that GML operates with no sensitivity parameter on the ratio of reinforcers obtained from the rich ( $r_R$ ) and lean ( $r_L$ ) schedules (see also Baum, 2002):

$$\frac{B_R}{B_L} = b_R \frac{r_R}{r_L}, \quad (3)$$

or, equivalently,

$$\log \frac{B_R}{B_L} = \log \frac{r_R}{r_L} + \log b_R, \quad (3')$$

where  $B_R$  and  $B_L$  are the rates of responding on the rich and lean schedules, respectively, and  $b_R$  is the bias towards the rich alternative. Note that the value of  $b_R$  can be greater than 1, smaller than 1, or equal to 1, so  $\log b_R$  can be either positive or negative or zero. In this formulation the matching function intercepts the ordinate at  $\log b_R$ . At this point,  $r_R = r_L$ , so the distinction between rich and lean is meaningless; if  $r_R$  is further reduced, however, the rich schedule to the right of the ordinate becomes the lean schedule, and vice versa. Thus, when Eq. (3') is plotted with the ratio of reinforcement rates ( $\log[r_1/r_2]$ ) on the abscissa, a discontinuity is observed at the ordinate.<sup>1</sup> Two possible discontinuities – step down or up – are described by the ideal matching functions in the top panels of Fig. 1; the step-

down pattern (left) corresponds to undermatching, whereas the step-up pattern corresponds to overmatching.

To the left of the ordinate in the top panels of Fig. 1, alternative 1 is lean and alternative 2 is rich; to the right of the ordinate, alternative 1 is rich and alternative 2 is lean. Tracing Eq. (3') from left to right, the function breaks at  $\log b'_{R=2}$ , which is  $\log b_R$  when alternative 2 is rich. The function then continues from  $\log b'_{R=1}$ , which is  $\log b_R$  when alternative 1 is rich. From Eq. (3), it may be shown that  $b'_{R=1} = -b'_{R=2}$ , as illustrated by the symmetry of the discontinuity around the origin (Fig. 1, top panels). A more general form of the model, which we call the fix-and-sample (FS) model, assumes a bias  $b_1$  towards alternative 1, such that

$$b_{R=1} = b_1 b'_{R=1}, \quad b_{R=2} = b_1 b'_{R=2}. \quad (4)$$

Coefficient  $b_1$  allows the matching function to be shifted up or down – not just symmetrically around the  $x$ -axis – to fit data. The bottom panels of Fig. 1 illustrate the fit of the FS model to data, based on the performance of one rat with two levels of effort for changing over alternatives (Aparicio, 2001). The FS model was implemented as Eq. (3') with two free parameters:  $b_R$  when  $r_1 < r_2$  (i.e.,  $b_{R=1}$ ) and  $b_R$  when  $r_1 > r_2$  (i.e.,  $b_{R=2}$ ). Log estimates of  $b_{R=1}$  and  $b_{R=2}$  were obtained using the method of least squares, and are shown in the bottom panels of Fig. 1; they can also be seen, respectively, as the  $y$ -intercepts of the solid lines on the right and left sides of each plot. Note that the discontinuities in the fitted functions are not vertically centered on zero, but on a negative number, indicating that  $b_1$  was negative—choice was biased toward alternative 2.

It is not obvious how a discontinuity in behavior between richer and leaner schedules can be incorporated to continuous-function local models of matching without violating their fundamental assumptions. The FS model (Eqs. (3') and (4)) describe global patterns of behavior allocation, but it does not explain why these equations must hold and does not specify local choice mechanisms whose aggregated operation would yield fix-and-sample patterns. We propose a hypothetical local choice mechanism that is consistent with the FS model. The mechanism is based on parsimonious assumptions of the internal state of an organism, which are specified by the formal relation between the system’s internal state, the environment’s influence, and the probability with which the system chooses each alternative.

## 2. A model of bipolar choice

This model was first introduced to explain some phenomena in human moral choice (Lefebvre, 1982, 1992). Consider the following situation. A person is facing an alternative: to tell the truth or to lie. Let the truth be “good” for this person and to lie “bad.” In addition, for telling the truth the person would receive \$10 and for lying \$10,000. This is an example of a situation in which a choice has two aspects, one moral and one utilitarian. In the moral aspect, the alternatives are bipolar. One of them can be called the positive pole and the other the negative pole. In the utilitarian aspect, the alternatives are assigned with numbers that correspond to their utility. In this particular example, the negative

<sup>1</sup> In Baum et al.’s (1999) model, matching is discontinuous at  $\log(B_R/B_L) = 0$  (preference indifference), and not necessarily when reinforcement rates are equal. This is because Baum et al. postulated their model as GML operating on the ratio of reinforcers obtained from the *preferred* and *nonpreferred* schedules. Eq. (3) diverges from Baum et al.’s to the extent that leaner schedules are preferred.

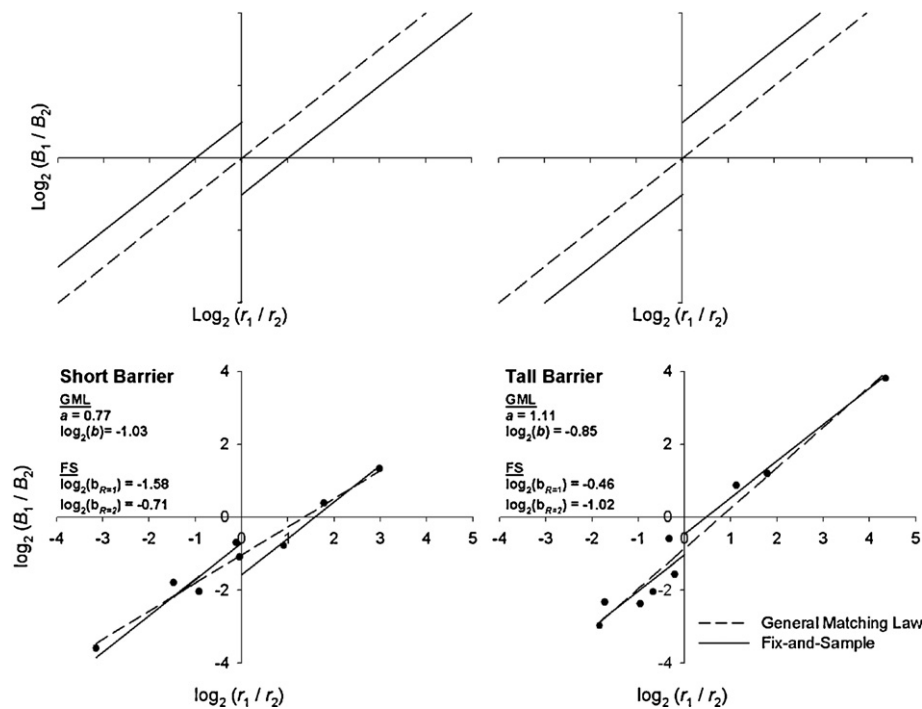


Fig. 1. (Top panels) Ideal patterns given by Eq. (3'). The pattern on the left panel corresponds to the case when  $\log b_R < 0$ , i.e., undermatching. The pattern on the right panel corresponds to the case when  $\log b_R > 0$ , i.e., overmatching. The dotted line corresponds to  $\log b_R = 0$ . (Bottom panels) Illustrative fitting of the Generalized Matching Law (GML; Eq. (2)) and the fix-and-sample model (FS; Eqs. (3') and (4)) to performance in concurrent variable-interval variable-interval (Conc VI VI) schedules of reinforcement. Relative rates of responding are shown as a function of relative rates of reinforcement in one rat (#62) when response levers were separated by a 30.5 cm barrier (left panel) and by a 45.7 cm barrier (adapted from Aparicio, 2001). The dashed line is the best fitting form of GML and the solid lines are the best fitting form of FS. Fitted parameters are displayed for each model.

pole (lie) is more profitable than the positive pole (truth). We cannot be certain of which alternative will be chosen, because polarity and utility are inconsistent with each other. This is a situation akin to the complex ambivalence self-control scenario described by Rachlin (2000).

The formal model of bipolar choice was initially constructed for predicting human choice in situations of this type; it may be represented in the following equation:

$$X = \frac{x_1}{x_1 + (1 - x_1)x_2}, \quad 0 < x_1 \leq 1, \quad 0 \leq x_2 \leq 1. \quad (5)$$

In this equation,  $X$  is the probability of choosing the positive pole,  $x_1$  is the relative utility of the positive pole, and  $x_2$  is a parameter characterizing the subject's inner state, the value of which is determined by a larger context. If  $x_2 = 0$  then  $X = 1$ , that is, choice is completely based on the polarity of the alternatives; if  $x_2 = 1$  then  $X = x_1$ , that is, choice is completely based on the utility of the alternatives. Thus,  $x_2$  is the relative decision weight given to local considerations of utility ("what pays more now?") and to global considerations of "goodness" ("what is better in the long run?").

After the model of bipolar choice had been constructed, it was found that it could also make predictions beyond the area of moral choice. For example, it could explain a few psychophysical phenomena, among them the non-linear relation between magnitude and categorical estimations of the same physical stimuli (length, weight, duration, area) (Lefebvre, 1992). The

model shed a new light on the asymmetry in evaluations given by people of their acquaintances in using constructs of the type strong-weak, fast-slow, etc. Experimental data have demonstrated that the frequency of choosing a positive adjective was equal to 0.62, not 0.5 as was expected; the model explained this shift (Adams-Webber, 1997; Lefebvre, 1980). These results led to the notion that the model expressed by Eq. (5) may describe not only human choice but animal choice as well.

We hypothesized that choices in Conc VI VI have two aspects, which may be called *utilitarian* and *positive-negative*. The utilitarian aspect relates to the immediate preferences of the animal at a local scale, and positive-negative aspects to the animal behavior in a larger, global, time scale. We can rewrite Eq. (5) as

$$\frac{1 - X}{X} = x_2 \frac{1 - x_1}{x_1}, \quad (6)$$

which allows us to see a parallel between the model of bipolar choice and Eq. (3). Lean and rich alternatives may be assigned with positive and negative polarity—the assignment criteria will be discussed further below. For illustration, let the lean alternative be the positive pole and the rich alternative the negative one:

$$X = \frac{B_L}{B_R + B_L}, \quad x_1 = \frac{r_L}{r_R + r_L}. \quad (7)$$

After substituting these values in Eq. (6) we obtain

$$\frac{B_R}{B_L} = x_2 \frac{r_R}{r_L}. \quad (8)$$

Eq. (8) is analogous to Eq. (3).

### 3. A new version of the model of bipolar choice

In this section, we demonstrate that Eq. (8) may be derived from a few assumptions regarding local choice. Let us first represent the behavior of an organism in a situation of bipolar choice as the following function:

$$X = \Phi(x_1, S), \quad 0 < X \leq 1, \quad 0 < x_1 \leq 1, \quad S \geq 0. \quad (9)$$

The values of variables  $X$  and  $x_1$  are interpreted as in the old model:  $X$  is the long-term probability of an organism choosing the positive alternative, and  $x_1$  is the instantaneous probability with which the local environment instigates the organism to choose the positive alternative.  $S$  is an internal variable that expresses the organism's preference for the positive alternative at a global scale. Let us assume that the local impact of reinforcement is already established (e.g., animals are predisposed to prefer more immediate rewards), whereas global preference must be acquired. We assume that with  $x_1$  being constant and  $S$  growing, the probability of choosing the positive pole,  $X$ , grows. If there is no long-term preference,  $S=0$ , the probability of choosing a pole is completely determined by local influence:

$$\Phi(x_1, 0) = x_1. \quad (10)$$

The probability with which the system chooses the positive alternative when the internal variable is equal to  $S$  will be designated as  $X_S$ . To find function  $\Phi(x_1, S)$ , we invoke one assumption, which we called the Axiom of Repeated Choice:

When the internal variable grows from  $S$  to  $S + \Delta S$  ( $0 < \Delta S < 1$ ;  $\Delta S$  is considered small) and  $x_1$  does not change, the procedure of choice is as follows. First, the system makes a preliminary choice with the probability of choosing the positive alternative equal to  $X_S$ . If the positive alternative is chosen, the system realizes its choice. If the negative alternative is chosen, then, with a small probability equal to  $\Delta S$ , the system *cancels* its choice and *repeats* the procedure of choice (with the probability of choosing the positive alternative equal to  $X_S$ ). The result of the repeated choice is realized no matter which alternative is chosen. (Lefebvre, 2004)

The empirical basis of the Axiom of Repeated Choice is incomplete, and thus it should be deemed as a hypothesis. Nonetheless, the phenomenon of repeated choice was observed in subjective estimation of stimulus intensity on a linear scale (Poulton et al., 1968). Subjects were presented with a stimulus and had to rate its intensity on a 1–100 scale. The experimenters found that after marking the scale, subjects often crossed it out and repeated the procedure of choice anew. Lefebvre (2006) demonstrated that these results could be analyzed in a scheme of choice between positive and negative alternatives.

The Axiom of Repeated Choice is depicted in Fig. 2. When there is no change in the internal variable  $S$ , choices are made

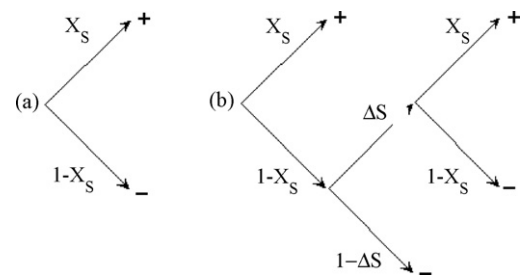


Fig. 2. Decision trees derived from the Axiom of Repeated Choice (a) when the value of the internal variable remains constant at  $S$ , and (b) when the value of the internal variable changes to  $S + \Delta S$ .

according to decision tree (a): the positive alternative is chosen with probability  $X_S$  and the negative alternative with probability  $1 - X_S$ . When  $S$  grows by  $\Delta S$ , decision tree (b) is applied. The initial choice in tree (b) is similar to the choice in tree (a), but if the negative alternative is chosen, there is a small probability,  $\Delta S$ , that the choice will be reconsidered.

It follows from tree (b) in Fig. 2 that the probability of choosing the positive alternative when the internal variable changes ( $X_{S+\Delta S}$ ) is the sum of the probability of choosing the positive alternative in the preliminary stage ( $X_S$ ) and of the joint probability of choosing the negative alternative in the preliminary stage ( $1 - X_S$ ), reconsidering the choice ( $\Delta S$ ), and finally choosing the positive alternative ( $X_S$ ):

$$X_{S+\Delta S} = X_S + (1 - X_S) \Delta S X_S. \quad (11)$$

As preference for the positive alternative at a global scale is acquired, the internal variable stabilizes and  $\Delta S \rightarrow 0$ . By considering  $X_S$  as a differentiable function with argument  $S$  and assuming  $\Delta S \rightarrow 0$ , we obtain the following differential equation (see Appendix A):

$$\frac{dX(S)}{dS} = (1 - X(S))X(S). \quad (12)$$

After solving Eq. (12) under condition  $X(0) = x_1$ , we find that

$$X = \Phi_{x_1}(S) = \frac{x_1}{x_1 + (1 - x_1) e^{-S}}. \quad (13)$$

Eq. (13) corresponds to Eq. (5) for bipolar choices, if  $x_2 = e^{-S}$ . Analogously to Eqs. (5) and (8), Eq. (13) can be transformed to

$$\frac{B^-}{B^+} = e^{-S} \frac{r^-}{r^+}, \quad (14)$$

where '+' corresponds to the positive alternative and '-' to the negative one. By finding the logarithms of the right and left sides of Eq. (14), we arrive to the final equation:

$$\ln \frac{B^-}{B^+} = \ln \frac{r^-}{r^+} - S. \quad (15)$$

This equation describes the relation between the frequencies of choosing positive and negative alternatives, the frequencies of their reinforcements, and the internal variable.

Let us now compare the Eq. (3') (from the FS model) and (15). If it is assumed that the logarithm base in Eq. (3') is  $e$  and  $\ln b_R = -S$ , and considering that  $S \geq 0$ , it follows that the richer

and leaner schedules are negative and positive, respectively. This is the scenario where matching is biased toward the leaner schedule and thus choices undermatch reinforcement. The polarity of the alternatives may be different, in which case  $\ln b_R = S$ , which implies that  $\ln b_R \geq 0$ . This is the scenario where matching is biased toward the richer schedule and thus choices overmatch reinforcement.

#### 4. The meaning of polarity

What does it mean for a lean or rich schedule to be “positive” or “negative”? In the original model of bipolar choice, polarity indicated the desirability of an alternative in a larger, ethical context. Prior research has shown that pigeons may become sensitive to global contingencies of reinforcement that are inconsistent with local contingencies, as when an alternative is preferred in the long run but not immediately (Heyman and Tanz, 1995; Sanabria et al., 2003). If polarity of alternatives signaled long-term contingencies, it would be expected that choices in Conc VI VI would be biased, relative to matching, in the direction that would optimize long-term reinforcement. Nonetheless, Houston and McNamara (1981) demonstrated that, with the exception of very extreme cases, overmatching optimizes rate of reinforcement in Conc VI VI, whereas most research reports undermatching patterns (Davison and McCarthy, 1988).

Another possibility is that matching is optimal at an evolutionary scale, and thus the positive polarity of lean schedules improves the Darwinian fitness of the undermatching organism. Indeed, the notion that matching is innate has been advanced (Gallistel et al., 2007). Because this hypothesis is hard to falsify, we will not speculate in candidate mechanisms for the natural selection of undermatching.

A third possibility involves the misattribution of reinforcement. If reinforcers are occasionally attributed to the wrong schedule, it would be expected that most misattributions would favor the lean schedule, because most reinforcers are provided by the rich schedule (Davison and Jenkins, 1985). Reinforcement misattribution would thus yield undermatching. It is possible that the polarity of alternatives reflects the attribution of reinforcement: Positive alternatives would be those that are attributed more reinforcement than they actually yielded.

The misattribution hypothesis makes informative, verifiable predictions of choice behavior. For instance, if a few assumptions are made on how reinforcement operates on behavior that precedes the effective response (Catania, 1971; Killeen, 1994), misattribution explains why delays to changeover between alternatives decrease undermatching (Boelens and Kop, 1983; Shull and Pliskoff, 1967). It cannot, however, explain why overmatching happens at all (Baum et al., 1999; but see Wearden, 1983).

In synthesis, there is no clear rule yet for assigning polarity to alternatives. Assuming undermatching as the modal pattern of choice in Conc VI VI, the model of bipolar choice predicts an initial tendency to match choices to reinforcement, and the progressive acquisition of a tendency to allocate more behavior to the lean (positive) schedule. This latter tendency may be driven by long-term contingencies, evolutionary predispositions, or by acquired errors in reinforcement attribution. Empirical research

is necessary to establish whether bias in matching is acquired and why.

#### 5. Prediction of local patterns of choice

We have demonstrated that the global pattern described by the FS model may be derived from a simple choice mechanism that learns to “doubt” before selecting one of the alternatives but not the other. To the extent that a choice procedure yields the global pattern predicted by the FS model, be it a discrete choice, concurrent schedules, or concurrent chain schedules of any kind, the bipolar choice model can provide a plausible local choice mechanism. Can the same mechanism describe local patterns of choice? The data reported by Baum et al. (1999) provide an informative constraint to any model of local choice: Visit durations to the preferred (rich) alternative, but not to the non-preferred (lean) alternative, covary with relative rate of reinforcement. The algorithm described in Fig. 2 would not yield this local pattern, mainly because it does not incorporate the burst-pause pattern of keypecking that yields a high autocorrelation of choices in concurrent schedules (Nevin and Baum, 1980). This system may also need to incorporate MacDonall’s (1999, 2000; MacDonall et al., 2006) insight that choice between concurrent schedules is constituted by choices between staying in and switching from each alternative. It is possible to conceive an undermatching organism that stays in the leaner alternative with probability  $X_S$  (Fig. 2), but that occasionally reconsiders its choice of staying in the richer alternative, with a probability of actually leaving ( $\Delta S$ ) that negatively covaries with the rich rate of reinforcement. Such organism would behave like Baum et al.’s pigeons.

Conceivability does not entail necessity, but then again, no model is strictly necessary. Ours is no exception, but it may be argued that, without further empirical constraints, there are too many degrees of freedom in the model to be of any use. This does not ring true, because we have shown that local patterns of choice actually falsify a strict version of the model. A non-strict version of the model could be the one in which a choice is made with the probability  $X_S$  only when Eq. (14) holds (Lefebvre, 2006); if this correlation is infringed, an organism begins restoring it by keeping the duration of visits to the lean alternative constant and increasing or decreasing the number of visits to the rich alternative as postulated in the bipolar model.

The model certainly needs to be developed to fully account for data but, more importantly, it also needs to make unique predictions that permit its empirical validation. Let us propose one prediction: If choice reconsideration is behaviorally expressed as orienting or moving toward the operandum without its actual activation, those actions should be evident in visits to one schedule – maybe at the end of response bursts – and not in visits to the other, and only during the acquisition phase. In Aparicio’s (2001) barrier choice paradigm, we expect that, when short barriers separate the alternatives and undermatching is observed, rats would occasionally move toward the lever in the rich schedule before climbing away, whereas the abandonment of the lean schedule would be more resolute. If the height of the barriers is raised and overmatching occurs, we expect that incomplete motions would be evident in the lean, not in the rich schedule.

Data on the topography of choice behavior is, unfortunately, scant. We hope that the model described here motivates further exploration.

### Acknowledgements

Federico Sanabria was supported by NIMH grant 1R01MH06686 to Peter R. Killeen. We thank William M. Baum, Peter R. Killeen, Nathalie Boutros, Douglas Elliffe, Alliston K. Reid, and two reviewers (one of them identified as Armando Machado) for their helpful comments on the draft of this paper, and Dr. Victorina Lefebvre without whose help this work would not be completed.

### Appendix A. The model of bipolar choice is a differentiable logistic function

Eq. (11) is an algebraic expression of the Axiom of Repeated Choice, where  $X_S$  is a given probability. Our next step is to find a differentiable function  $X(S)$  that corresponds to Eq. (11) and where  $X(0) = x_1$ . The latter constraint signifies the exclusive determination of choice by local contingencies when the internal variable is zero.

We do not know in advance if, among an infinite number of differentiable functions, there are such that can be approximated by Eq. (11) with an infinitely small value in comparison with  $\Delta S$ . This infinitely small value is designated  $o(\Delta S)$  and defined as  $o(\Delta S)/\Delta S \rightarrow 0$ , when  $\Delta S \rightarrow 0$ . If there is at least one function of the type we search for, then for the small increase of  $S$  we can write:

$$X(S + \Delta S) = X(S) + (1 - X(S))X(S) \Delta S + o(\Delta S) \quad (\text{A.1})$$

or

$$\frac{\Delta X(S)}{\Delta S} = (1 - X(S))X(S) + \frac{o(\Delta S)}{\Delta S}. \quad (\text{A.2})$$

For the limit at  $\Delta S \rightarrow 0$ , we obtain differential equation (12) and solve it for  $X(0) = x_1$ . As a result, we obtain Eq. (13), which is a special case of the model of bipolar choice, and confirm with a direct test that it satisfies Eq. (A.1).

### References

- Adams-Webber, J., 1997. Self-reflexion in evaluating others. *Am. J. Psychol.* 110, 527–541.
- Aparicio, C.F., 2001. Overmatching in rats: the barrier choice paradigm. *J. Exp. Anal. Behav.* 75, 93–106.
- Aparicio, C.F., Baum, W.M., 2006. Fix and sample with rats and the dynamics of choice. *J. Exp. Anal. Behav.* 86, 43–63.
- Baum, W.M., 1974. On two types of deviation from the matching law: bias and undermatching. *J. Exp. Anal. Behav.* 22, 231–242.
- Baum, W.M., 2002. From molecular to molar: a paradigm shift in behavior analysis. *J. Exp. Anal. Behav.* 78, 95–116.
- Baum, W.M., Rachlin, H., 1969. Choice as time allocation. *J. Exp. Anal. Behav.* 12, 861–874.
- Baum, W.M., Schwendiman, J.W., Bell, K.E., 1999. Choice, contingency discrimination, and foraging theory. *J. Exp. Anal. Behav.* 71, 355–373.
- Boelens, H., Kop, P.F.M., 1983. Concurrent schedules—spatial separation of response alternatives. *J. Exp. Anal. Behav.* 40, 35–45.
- Catania, A.C., 1971. Reinforcement schedules: the role of responses preceding the one that produces the reinforcer. *J. Exp. Anal. Behav.* 15, 271–287.
- Davison, M.C., Jenkins, P.E., 1985. Stimulus discriminability, contingency discriminability, and schedule performance. *Anim. Learn. Behav.* 13, 77–84.
- Davison, M.C., McCarthy, D., 1988. *The Matching Law: A Research Review*. Lawrence Erlbaum Associates, New Jersey.
- De Villiers, P.A., Herrnstein, R.J., 1976. Toward a law of response strength. *Psychol. Bull.* 83, 1131–1153.
- Gallistel, C.R., King, A.P., Gottlieb, D., Balci, F., Papachristos, E.B., Szalecki, M., Carbone, K.S., 2007. Is matching innate? *J. Exp. Anal. Behav.* 87, 161–199.
- Herrnstein, R.J., 1961. Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* 4, 267–272.
- Herrnstein, R.J., 1982. Melioration as behavioral dynamism. In: Commons, M.L., Herrnstein, R.J., Rachlin, H. (Eds.), *Quantitative Analyses of Behavior. Matching and Maximizing Accounts*, vol. 2. Ballinger, Cambridge, MA, pp. 433–458.
- Heyman, G.M., Tanz, L., 1995. How to teach a pigeon to maximize overall reinforcement rate. *J. Exp. Anal. Behav.* 64, 277–297.
- Houston, A.I., McNamara, J., 1981. How to maximize reward rate on two variable-interval paradigms. *J. Exp. Anal. Behav.* 35, 367–396.
- Killeen, P.R., 1994. Mathematical principles of reinforcement. *Behav. Brain Sci.* 17, 105–172.
- Lau, B., Glimcher, P.W., 2005. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555–579.
- Lefebvre, V.A., 1980. An algebraic model of ethical cognition. *J. Math. Psychol.* 22, 83–120.
- Lefebvre, V.A., 1982. *Algebra of Conscience*. D. Reidel, Dordrecht, Holland.
- Lefebvre, V.A., 1992. *A Psychological Theory of Bipolarity and Reflexivity*. The Edwin Mellen Press, Lewiston, NY.
- Lefebvre, V.A., 2004. Bipolarity choice and entro-field. In: *Proceedings of the Eighth World Multi-Conference on Systemics, Cybernetics and Informatics*, vol. IV, pp. 95–99.
- Lefebvre, V.A., 2006. *Research on Bipolarity and Reflexivity*. The Edwin Mellen Press, Lewiston, NY.
- Lobb, B., Davison, M.C., 1975. Performance in concurrent interval schedules: a systematic replication. *J. Exp. Anal. Behav.* 24, 191–197.
- MacDonall, J.S., 1999. A local model of concurrent performance. *J. Exp. Anal. Behav.* 71, 57–74.
- MacDonall, J.S., 2000. Synthesizing concurrent interval performances. *J. Exp. Anal. Behav.* 74, 189–206.
- MacDonall, J.S., Goodell, J., Juliano, A., 2006. Momentary maximizing and optimal foraging theories of performance on concurrent VR schedules. *Behav. Process.* 72, 283–299.
- Nevin, J.A., Baum, W.M., 1980. Feedback functions for variable-interval reinforcement. *J. Exp. Anal. Behav.* 34, 207–217.
- Poulton, E.C., Simmonds, D.C.V., Warren, R.M., 1968. Response bias in very first judgments of the reflectance of grays: numerical versus linear estimates. *Percept. Psychophys.* 3, 112–114.
- Rachlin, H., 2000. *The Science of Self-Control*. Harvard University Press, Cambridge, MA.
- Sanabria, F., Baker, F., Rachlin, H., 2003. Learning by pigeons playing against tit-for-tat in an operant prisoner's dilemma. *Learn. Behav.* 31, 318–331.
- Shull, R.L., Pliskoff, S.S., 1967. Changeover delay and concurrent schedules—some effects on relative performance measures. *J. Exp. Anal. Behav.* 10, 517–527.
- Staddon, J.E.R., Motheral, S., 1978. On matching and maximizing in operant choice experiments. *Psychol. Rev.* 85, 436–444.
- Wearden, J.H., 1983. Undermatching and overmatching as deviations from the matching law. *J. Exp. Anal. Behav.* 40, 333–340.