Beyond Matching: Behavioral Probability Currents in Multi-Alternative Choice

A theoretical extension of matching-based changeover models with numerical tests of phase-driven circulation

Theoretical and simulation paper

Abstract

The matching law and its generalized form characterize how behavior is allocated among concurrently available alternatives as a function of reinforcement. Recent changeover theory has shown that two-alternative concurrent performance can also be represented as a continuous-time Markov chain whose dynamics are completely determined by the leaving rate from each alternative. The present paper asks what changes when the choice set contains a loop: three or more alternatives. We show that stationary allocation no longer determines the transition process. A three-state system can preserve exactly the same allocation while supporting an independent, divergence-free probability current around the choice network. We call this quantity behavioral current. We then examine a minimal periodically driven Markov model in which state-specific leaving readiness and symmetric pairwise accessibility vary over time. Every frozen environmental configuration obeys detailed balance and has zero stationary current, yet cyclic modulation generates a nonzero integrated current. Numerical integration shows that current reverses when the phase relation is reversed, is approximately sinusoidal in phase ($R^2 = 0.999989$ for the tested parameters), vanishes under one-factor controls, and approaches a finite value per cycle in the slow-driving regime while current per unit time declines. Mean state allocation remains nearly unchanged across opposite-current conditions. These results identify a theoretically independent dimension of multi-alternative behavior that matching ratios alone cannot recover. The proposal is positioned as an extension of matching and stay/switch theory rather than a replacement, and it yields direct experimental tests for human and nonhuman concurrent choice.

1. Introduction

Herrnstein's matching relation established that relative responding on concurrently available alternatives tracks relative obtained reinforcement. Baum's generalized matching law subsequently represented systematic sensitivity and bias using a log-linear relation. These molar formulations have been exceptionally productive because they compress a stream of behavior into relative allocation. That compression, however, necessarily discards information about the order in which alternatives are visited.

Behavior-analytic theories have long recognized that switching matters. The stay/switch model treats choice from the organism's current location: at an alternative, the relevant action is to

remain or to leave. MacDonall showed that reinforcers for staying and switching can account for patterns not captured by a standard generalized matching analysis. More recently, Avellaneda (2025) formalized two-alternative changeover behavior using a continuous-time Markov chain. In the two-alternative case, the leaving rate from each alternative suffices to determine the entire chain. This result is powerful, but its restriction to two alternatives is mathematically consequential.

With three alternatives, leaving a state no longer specifies the destination. The transition network contains a loop, and a new degree of freedom becomes possible: net circulation around that loop. Multi-alternative matching has been studied empirically, including three- and four-alternative preparations, but the usual dependent variables remain allocations of responses, time, entries, or reinforcement. The present paper isolates a complementary quantity: antisymmetric transition flow, or probability current.

The central claim is modest but testable: matching can correctly describe where behavior is allocated while failing to describe how behavior flows among alternatives. For networks with loops, a complete dynamic description therefore requires both occupation variables and current variables.

2. From Matching to Current

2.1 Allocation

For two alternatives, the generalized matching law can be written as

$$B_1/B_2 = b\,(R_1/R_2)^s$$

where B denotes behavior allocation, R denotes obtained reinforcement, b is bias, and s is sensitivity. A multi-alternative extension can be expressed in terms of stationary occupation probabilities p_i. These quantities describe how much of the behavioral stream resides in each alternative.

2.2 Transition flows

Let q_{ij} denote the instantaneous transition rate from alternative i to alternative j. The stationary one-way flow is

$$F_{ij} = p_i\, q_{ij}.$$

The antisymmetric component of this flow is

$$J_{ij} = F_{ij} - F_{ji} = p_i q_{ij} - p_j q_{ji}.$$

At stationarity, probability conservation requires the current entering each state to equal the current leaving it. In a fully connected three-state system this implies a single independent cycle current:

$$J = J_{AB} = J_{BC} = J_{CA}.$$

A positive value denotes net circulation A→B→C→A; a negative value denotes the reverse direction. Crucially, the occupation vector p does not determine J.

2.3 Identifiability result

Consider a three-state continuous-time Markov chain with clockwise transition rates $1+c$ and counterclockwise transition rates $1-c$, where $|c|<1$. For every c, the stationary occupation is exactly

$$p_A = p_B = p_C = 1/3,$$

but the cycle current is

$$J = 2c/3.$$

Thus infinitely many transition processes can have identical matching allocation and different directional organization. This is not an empirical claim about organisms; it is an identifiability theorem for the behavioral description. Any theory based only on stationary allocation is unable, in principle, to distinguish these processes.

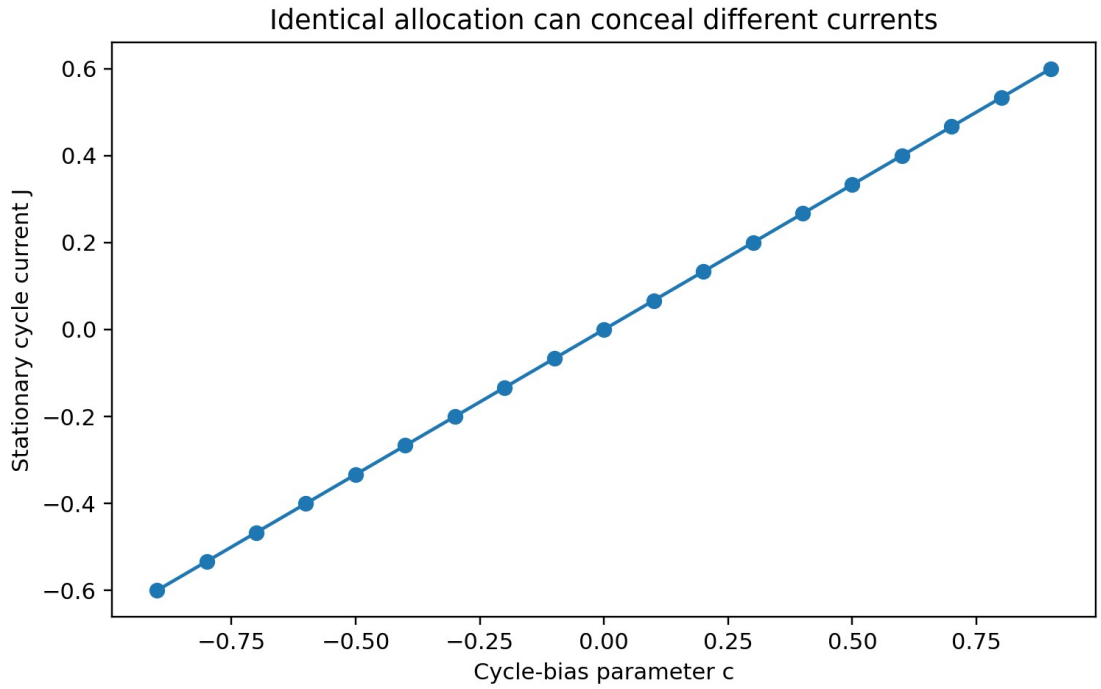

Figure 1. Exact allocation-current dissociation. Stationary allocation remains 1/3 at every alternative while cycle current changes linearly with the hidden transition-bias parameter c.

3. Phase-Driven Current Model

We next ask whether a current can be generated without building a directional preference into any frozen environmental state. Let the transition rate factor into a state-specific leaving-readiness term $r_i(t)$ and a symmetric accessibility term $g_{ij}(t)=g_{ji}(t)$:

$$\mathbf{q_{ij}(t) = r_i(t)\, g_{ij}(t).}$$

For frozen parameters, the stationary distribution is proportional to $1/r_i$, and detailed balance holds on every edge. Consequently, every static configuration has zero stationary current. This factorization is a modeling assumption, not a claim that all organismic switching obeys it.

For the minimal three-state model, only r_A and g_AB were modulated:

$$\mathbf{r_A(t)=exp[a\ cos(2\pi t/T)]}$$

$$\mathbf{g_AB(t)=exp[b\ cos(2\pi t/T + \varphi)]}$$

All other r and g terms were fixed at 1. Parameter a controls the amplitude of leaving-readiness modulation, b controls accessibility modulation, T is the environmental period, and φ is their phase offset. Reversing φ reverses the orientation of the path traced through parameter space.

4. Simulation Methods

The master equation dp/dt = pQ(t) was integrated numerically with a fourth-order Runge-Kutta scheme. Each environmental cycle was discretized into 1,000-1,400 steps depending on the analysis. Cycles were iterated until the start-of-cycle occupation vector converged to its periodic steady state; the residual start/end difference was at machine precision in the reported runs. The integrated edge currents were then calculated over one complete cycle and averaged across the three edges.

Primary simulations used a=b=0.5. Four analyses were conducted: (a) a phase sweep from −180° to +180° at T=200; (b) a speed sweep from T=5 to T=640 at φ=+90°; (c) forward, reverse, retention-only, accessibility-only, in-phase, and anti-phase controls at T=320; and (d) an exact stationary construction demonstrating identical occupation with variable current. A separate Gillespie simulation of the stationary construction confirmed the analytic current within Monte Carlo error.

The simulation is a proof of model behavior. No human or nonhuman participant data were used.

5. Results

5.1 Phase controls direction

The phase sweep was almost perfectly described by a sinusoid. Least-squares fitting yielded $Q(\varphi) = 0.02849 \sin\varphi + 0.00031 \cos\varphi + 0.000001$, $R^2 = 0.999989$. Thus the sign of circulation was controlled by the temporal ordering of the two environmental modulations rather than by a persistent directional bias.

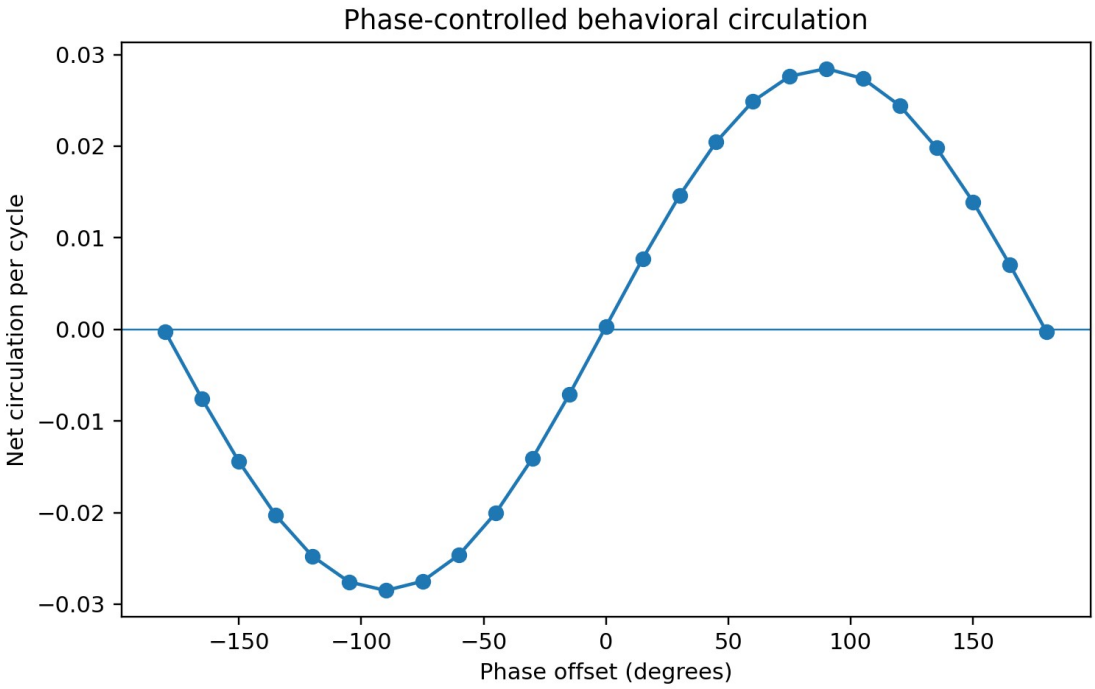

Figure 2. Net circulation as a function of phase offset. Positive and negative phase relations produce opposite directional currents.

5.2 Time reversal reverses current while allocation is nearly preserved

At T=320, the +90° condition produced Q=0.02848, whereas the −90° condition produced Q=-0.02850. Mean allocations were A=0.3378, B=0.3312, C=0.3311 in the forward condition and A=0.3379, B=0.3310, C=0.3311 in the reverse condition. The largest deviation of any mean allocation from 1/3 was 0.0046. Opposite transition organization therefore occurred with nearly indistinguishable molar allocation.

Condition	Q/cycle	pA	pB	pC
forward_+90	0.02848	0.3378	0.3312	0.3311
reverse_-90	-0.02850	0.3379	0.3310	0.3311
retention_only	0.00000	0.3378	0.3311	0.3311
access_only	0.00000	0.3333	0.3333	0.3333
in_phase	0.00021	0.3378	0.3311	0.3311
anti_phase	-0.00018	0.3378	0.3311	0.3311

Table 1. Periodic steady-state controls at T=320. Retention-only and accessibility-only manipulations do not generate cycle current; in-phase and anti-phase currents shrink toward zero in the slow-driving limit.

5.3 One-factor controls

Varying leaving readiness alone produced Q=3.97e-18; varying accessibility alone produced Q=0.00e+00. The result demonstrates that current in this model is generated by the coordinated temporal relationship between the two parameter classes, not by either modulation in isolation.

5.4 Slow-driving behavior

As T increased from 5 to 640, circulation per cycle approached a plateau, increasing from 0.02352 to 0.02851. In contrast, circulation per unit time fell from 0.004703 to 0.000045. This is the expected qualitative signature of a geometric pumping contribution: slowing the environmental traversal preserves an approximately finite transported amount per cycle while reducing the transport rate per unit time.

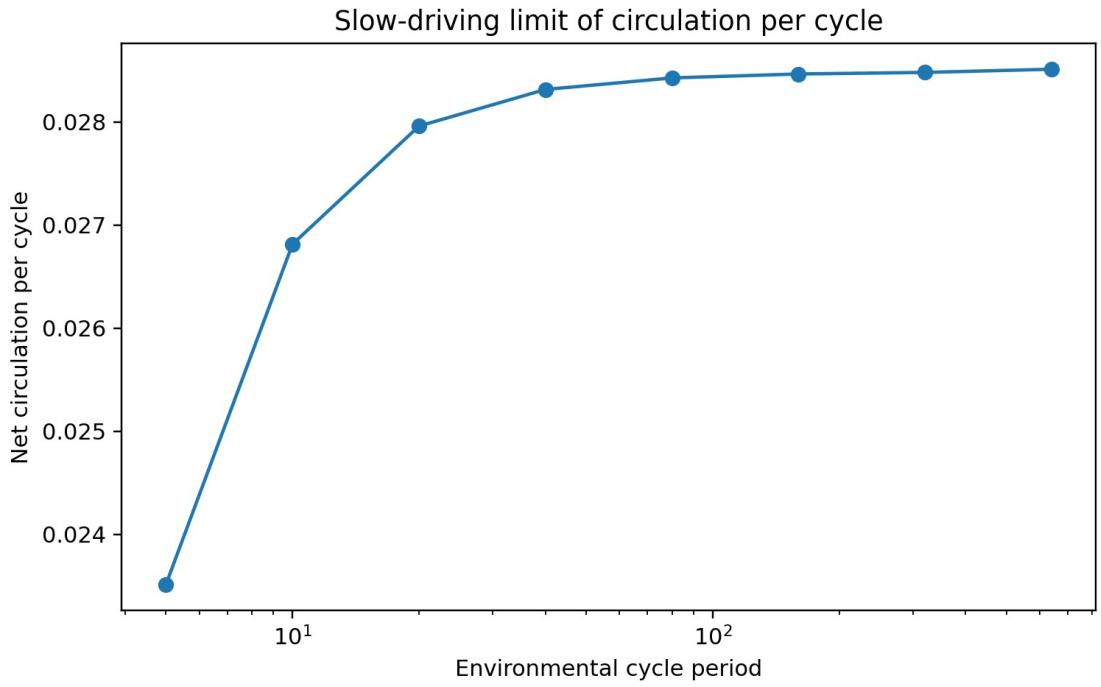

Figure 3. Integrated circulation per environmental cycle approaches a finite slow-driving value.

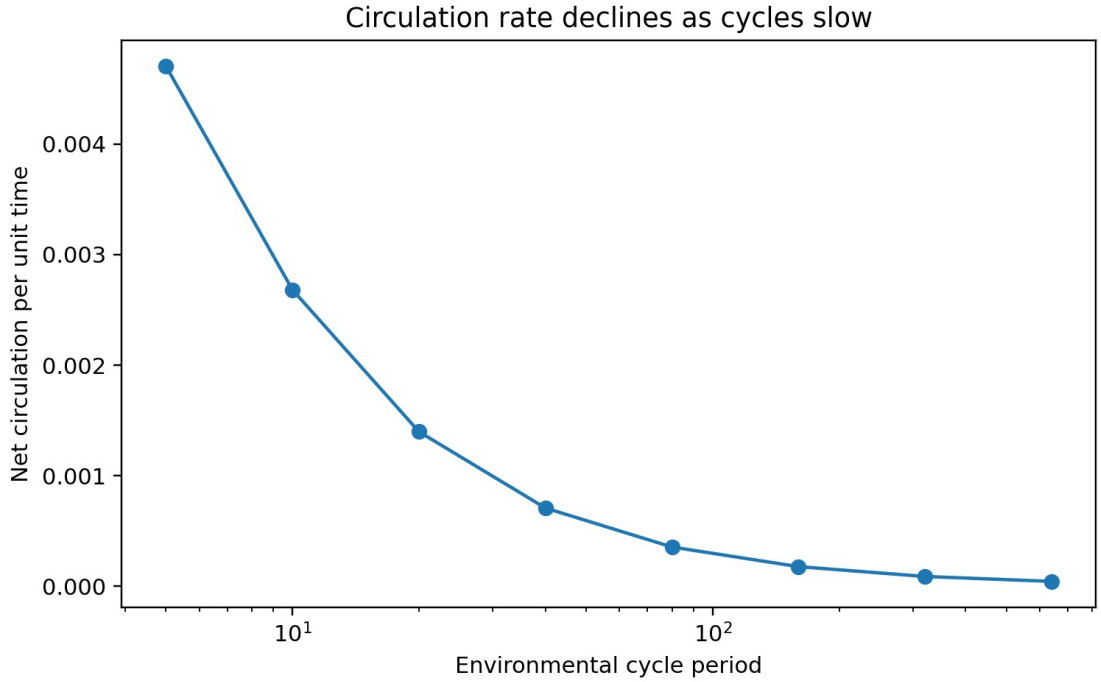

Figure 4. Circulation per unit time decreases as the environmental cycle is slowed.

6. Relation to Existing Choice Theory

6.1 Matching is retained as an allocation law

Nothing in the present formulation requires rejecting matching. Matching ratios can remain excellent descriptors of occupation while current variables describe additional temporal organization. The proposal is therefore a decomposition, not a replacement: allocation specifies where behavior resides; current specifies net directional flow through the network.

6.2 Why the distinction is absent in the two-alternative case

Avellaneda's two-alternative Markov formulation is complete because a departure from one alternative uniquely determines the destination. A two-node choice graph has no independent cycle. Once there are three alternatives, the leaving rate from A specifies only the total departure hazard; it does not determine how departures divide between B and C. Branching probabilities introduce degrees of freedom that stationary allocation alone does not identify.

6.3 Relation to stay/switch theory

The stay/switch model already makes the organism's current state and its switching contingencies central. The proposed extension adds destination structure. With two alternatives, 'switch from A' implies B. With three alternatives, 'switch from A' decomposes into at least two destination-specific transitions. Behavioral current is the antisymmetric component of those destination-specific flows.

6.4 Multi-alternative matching

Three- and four-alternative matching preparations show that allocation principles generalize beyond binary choice. The present point is orthogonal: even a perfect fit to multi-alternative allocation cannot identify all trajectory statistics. The missing information appears whenever the choice graph contains loops.

7. General Network Formulation

For a connected behavioral network with V states and E undirected connections, the dimension of its cycle space is

$$\mathbf{L = E - V + 1.}$$

L gives the number of independent loop currents that may exist after stationarity constraints are imposed. A binary choice has V=2, E=1, and L=0. A fully connected three-choice system has V=3, E=3, and L=1. A fully connected four-choice system has V=4, E=6, and L=3. Thus the amount of dynamic information omitted by pure allocation measures increases as the topology of the choice environment becomes richer.

A compact descriptive framework is therefore

Behavioral state = occupations $\{p_i\}$ + independent cycle currents $\{J_\gamma\}$.

This 'matching-current decomposition' is a proposed descriptive extension. Whether the current terms themselves obey simple reinforcement-based laws is an empirical question.

8. Experimental Predictions

A direct experiment requires at least three continuously available alternatives and complete recording of transitions. The strongest test would independently calibrate manipulations intended to affect leaving readiness and pairwise switching accessibility, then impose cyclic schedules while maintaining approximately matched long-run reinforcement totals.

- Time-reversal prediction. Reversing the environmental modulation should reverse the sign of the cycle current while leaving molar allocation approximately unchanged.
- Phase-response prediction. For small, slow modulations, current should vary approximately sinusoidally with phase offset.
- One-factor controls. Modulating only leaving readiness or only symmetric accessibility should not sustain a geometric cycle current under the minimal model.
- Speed dissociation. In the slow-driving regime, current per environmental cycle should approach a finite value while current per unit time declines as the period lengthens.
- Persistence test. If the modulation is stopped and the environment is held fixed, sustained current should disappear in the minimal model. Persistent circulation would imply learned directional structure or another mechanism.

A practical estimator for a three-alternative cycle is

$$\hat{J} = [N_{AB} + N_{BC} + N_{CA} - N_{BA} - N_{CB} - N_{AC}] / (3T_{obs}).$$

The critical comparison is not whether a flexible reinforcement-learning model can retrospectively fit the data, but whether competing models predict unseen phase, speed, and time-reversal conditions from parameters estimated independently.

9. Limitations and Novelty Boundary

The simulation establishes a mathematical possibility, not an organismic mechanism. The factorization $q_{ij}=r_i g_{ij}$ is deliberately minimal and may fail when reinforcement, effort, attention, discriminability, or history simultaneously affect both leaving and destination selection. Experimental manipulations must therefore be calibrated rather than assigned psychological labels by fiat.

Probability currents and stochastic pumping are established concepts in nonequilibrium statistical mechanics. Markov descriptions of choice, switching analyses, stay/switch theory, and multi-alternative matching are also established. The candidate contribution is narrower: treating divergence-free cycle current as an explicitly behavioral dependent variable in multi-alternative concurrent choice, showing why it is formally unidentifiable from matching allocation, and

deriving a phase-reversal experimental signature that can manipulate current with little change in allocation. A comprehensive priority claim would require a systematic literature review beyond the targeted search used here.

The present work also does not imply that recurrent real-world behavior is generally externally 'pumped.' Learning, conditioned reinforcement, discriminative control, response effort, verbal behavior, and longer-timescale contingencies remain plausible and often more appropriate explanations. The current framework is valuable only where it adds discriminating predictions.

10. Discussion

The central result is that an allocation law and a trajectory law answer different questions. Matching summarizes the fraction of behavior associated with alternatives. A probability current summarizes directional asymmetry in transitions among those alternatives. Binary choice collapses these descriptions because its graph contains no loop; multi-alternative choice separates them.

That separation suggests a new measurement target for behavioral science. Two organisms may match reinforcement identically and spend the same proportion of time at every alternative, yet traverse those alternatives in opposite directions. Conversely, an intervention may leave allocation almost unchanged while reorganizing the behavioral trajectory. Such changes would be invisible to a purely molar allocation measure.

The phase-driven simulation makes the distinction experimentally sharp. No frozen environmental state in the model contains a preferred circulation direction. Direction appears only from the temporal ordering of otherwise reversible environmental configurations. Reversing that order reverses the current. The result therefore offers a stringent empirical test rather than a post hoc narrative.

If supported experimentally, the matching-current decomposition would extend quantitative choice theory from distributions over alternatives to flows through alternative networks. The most conservative interpretation is not that matching is incomplete in its intended domain, but that multi-alternative behavior contains additional information that allocation was never designed to represent.

11. Conclusion

The generalized matching law can remain correct about behavioral allocation while being silent about directional transition flow. In a three-alternative system, an independent cycle current can vary without any change in stationary allocation. A minimal periodically driven Markov model further shows that environmental timing can generate and reverse such current even though each frozen condition is non-circulating. These results motivate a matching-current decomposition of multi-alternative behavior and a direct experimental program based on transition counts, time reversal, phase sweeps, and one-factor controls.

References

Avellaneda, M. A. (2025). A model of changeover behavior in two-alternative choice. Journal of the Experimental Analysis of Behavior, 124(1), e70025. https://doi.org/10.1002/jeab.70025

Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. Journal of the Experimental Analysis of Behavior, 22(1), 231-242. https://doi.org/10.1901/jeab.1974.22-231

Beeby, E., & Alsop, B. (2017). Choosing among multiple alternatives: Relative and overall reinforcer rates. Journal of the Experimental Analysis of Behavior, 108(2), 204-222. https://doi.org/10.1002/jeab.269

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. Journal of the Experimental Analysis of Behavior, 4(3), 267-272. https://doi.org/10.1901/jeab.1961.4-267

Kangas, B. D., Berry, M. S., Cassidy, R. N., Dallery, J., Vaidya, M., & Hackenberg, T. D. (2009). Concurrent performance in a three-alternative choice situation: Response allocation in a Rock/Paper/Scissors game. Behavioural Processes, 82(2), 164-172. https://doi.org/10.1016/j.beproc.2009.06.004

MacDonall, J. S. (2009). The stay/switch model of concurrent choice. Journal of the Experimental Analysis of Behavior, 91(1), 21-39. https://doi.org/10.1901/jeab.2009.91-21

Rahav, S., Horowitz, J., & Jarzynski, C. (2008). Directed flow in nonadiabatic stochastic pumps. Physical Review Letters, 101, 140602. https://doi.org/10.1103/PhysRevLett.101.140602

Reproducibility Materials

The accompanying simulation bundle contains the Python source, phase sweep, speed sweep, control results, exact identifiability construction, Monte Carlo check, and summary JSON used to generate the reported figures and numerical values.