Self-Triggerable Reward and Motor Resonance: A Dynamical Account of Attraction Without Specified Targets

Flyxion
Independent Researcher

November 2026

Abstract

Accounts of attraction usually treat the objects that attract as fixed by selection, either directly, as templates for reproductively relevant features, or indirectly, as consequences of the biases of the receiver's sensory system. This paper develops a third class of account in which the reward circuit is reachable from internal states alone, so that reward can be generated without any external object, and external objects acquire value through the trajectories they induce in the observer's own sensorimotor dynamics. The account is formalized with a state space for the body, a reward functional on trajectories, a percept map into that state space, and separate functionals for resonance and for reward. Four groups of results are given. Trajectory equivalence of objects refines reward equivalence, and value factors through the observer's induced trajectory. Classes of value function that do not depend on the observer's motor parameters are invariant under intervention on those parameters, whereas an explicit frequency-tuning law derived from the Adler equation shows that a resonance account is not. Under replicator dynamics on preference weights, satiation determines whether preference concentrates on one resonance class or remains plural. A comparison-lemma bound gives a persistence regime in which a self-sustained rewarding attractor survives any admissible outcome signal, and a contraction bound limits the outcome sensitivity of observable action, which together define a sphexish regime that a slow reinforcement law for the coupling strength drives the system toward. The paper states what each class of account forbids, specifies interventions that would discriminate among them, and states the conditions under which they cannot be distinguished. It does not claim that selection is irrelevant to attraction, only that the evolution of a reward system does not entail that the objects capable of activating it are evolutionarily specified.

Keywords: reward, motor resonance, attraction, receiver bias, synchronization, replicator dynamics, stereotyped behavior, identifiability

1 Introduction

The standard explanation of why organisms are attracted to particular objects proceeds by way of the biological significance of the object. A sexually attractive appearance is said to engage mechanisms that evolved because responding to that appearance was associated with reproductive success, and the experience of attraction is explained by pointing to what the stimulus signified for the lineage. A more recent and more careful family of accounts, the sensory-bias literature, relaxes this by allowing that preferences can arise from pre-existing properties of the receiver's perceptual system, so that selection need not specify every attractive target directly [21, 9]. Both families, in the forms formalized here, agree on one structural point, which is that the value of an object to an observer is determined by features of the object, filtered either through an evolved template or through a sensory tuning curve. The observer's own motor organization plays no role in determining which objects are valued. Sensory-bias theory in general is not committed to this, since sensory and motor systems are coupled, and the point is taken up in Section 8.

This paper examines the consequences of denying that structural point. The central hypothesis is that the reward system is reachable from internal states alone, so that reward does not require an external object, and that perceiving certain external dynamics induces corresponding activity in the perceiver's own sensorimotor system, some of which approaches the set of trajectories from which reward is accessible. Under this hypothesis an object attracts to the extent that its induced trajectory in the observer approaches the reward-accessible set of the observer's own sensorimotor dynamics. The object is a proxy for an internally generated experience, and which objects can serve as proxies depends on the observer's repertoire of motor dynamics and on the history of reinforcement, not only on the features of the object. The hypothesis is meant to separate three matters that evolutionary explanations sometimes compress into a single story: what an organism attends to, what makes an experience rewarding, and why the capacity for that experience exists at all.

The hypothesis is deliberately not a rejection of evolutionary biology. A reward circuit, a perceptual system, and a motor system are evolved structures, and the coupling between them may itself be an adaptation. The claim at issue is narrower. It is that the existence of an evolved reward system does not entail that the objects capable of activating it are themselves evolutionarily specified. In discussions of the evolution of agency, evolved dispositions are understood to bias behavior without fully determining it [25]. The present question is complementary: it concerns how the objects of such dispositions are fixed, and whether selection must do the fixing or whether it can instead produce a dynamical system whose objects of attraction emerge through perception, resonance, reinforcement, and, in some regimes, rigid closure.

A second motivation is the phenomenon that the literature calls sphexishness, the persistence of a behavioral routine after it has ceased to serve the purpose for which it appears to exist. The term was coined by Hofstadter, in a column of 1982 collected in 1985, and taken up by Dennett, both drawing on an anecdote about the digger wasp that originates with Wooldridge [27, 11, 7]. Keijzer has shown that the anecdote is repeated in the cognitive sciences far beyond what the ethological evidence supports, so that the rigidity of the wasp

is better treated as a tendency than as an absolute [14]. Whatever the status of the wasp, the formal question remains: under what conditions does a system whose reward is generated by its own activity become insensitive to the success or failure of that activity in the world? The resonance account gives a specific answer, developed in Section 6, and the answer links preference formation, which is usually discussed under learning, to behavioral rigidity, which is usually discussed under instinct.

The paper proceeds as follows. Section 2 fixes the primitives and states the hypotheses separately, because the argument is weakened when they are bundled. Section 3 distinguishes resonance from reward and defines proxy equivalence. Section 4 compares the resonance class with two motor-independent classes, a template class and a receiver-bias class, and derives an intervention under which they differ. Section 5 treats the formation of preference under reinforcement. Section 6 treats closure and sphexishness. Section 7 turns to identifiability and to the interventions that would discriminate among the classes. Section 8 states the relation to selection, Section 9 extends the framework to rhythm, sound, and gait, and Section 10 states limitations. The mathematics is deliberately elementary, since the aim is a precise statement of what the hypothesis says and forbids, not a demonstration that it is true.

2 Primitives and Self-Triggerability

Let $x \in X \subseteq \mathbb{R}^n$ denote the internal sensorimotor and physiological state of an organism i. The state evolves according to

$$\dot{x} = F(x, u; \theta_i), \tag{1}$$

where $u \in U$ is an input and $\theta_i = (\theta_i^m, \theta_i^r)$ collects the parameters that describe the organism, divided into motor and dynamical parameters θ_i^m, which determine what trajectories the body can produce, and reward parameters θ_i^r, which determine how trajectories are valued. The input set is divided into internal controls U_{int}, which the organism generates with no external object present, and external inputs, which depend on objects in the environment. The reward accrued along a trajectory $x(\cdot)$ of duration T is

$$\mathcal{R}_T[x] = \int_0^T r\big(x(t), \dot{x}(t); \theta_i^r\big)\, \mathrm{d}t, \tag{2}$$

where r is an instantaneous reward rate. Whether r is read as hedonic impact, incentive salience, or a learning signal is a modeling choice that the literature on dopamine shows is not innocent [2, 1]. The results below do not depend on the reading, but any application must state which is intended, because repeated pursuit is not by itself evidence of pleasure.

Definition 1 (Self-triggerability)**.** Organism i is *self-triggerable* if there exist an initial state x_0, an internal control $u \in U_{\text{int}}$, and a baseline ρ (for instance the reward accrued along the rest trajectory) such that the solution x_u of (1) satisfies $\mathcal{R}_T[x_u] > \rho$.

Hypothesis 1 (Self-triggerability)**.** *Organisms with the reward systems at issue are self-triggerable.*

Hypothesis 1 is the premise from which the rest of the argument departs. It is an existence statement and is the most empirically secure of the four hypotheses. Intracranial self-stimulation established that reward can be obtained with no external object at all [16], and the comparative literature on primate sexuality documents self-directed sexual behavior across many species of primate [8]. In the vocabulary of Section 2 these observations say that the reward-accessible set can be reached by trajectories generated entirely from within. A consequence worth stating explicitly is that, for a self-triggerable organism, the pleasure of a particular kind of activity is discoverable independently of any mate, offspring, or other biologically significant object. The activity can be found by the organism's own movement, and its value is established before any object is attached to it.

Hypothesis 1 does not establish more than it states. It does not establish that the organism can initiate the rewarding trajectory voluntarily at any time, that every trajectory induced from outside is rewarding, or that the trajectories recruited when an external object attracts are the ones reached by self-stimulation. Those further claims are separate hypotheses, and keeping them separate is what allows the theory to be tested piece by piece instead of being defended as a whole.

Hypothesis 2 (Induced overlap)**.** *Observing an external object o induces a trajectory in the observer's state space, and for some objects this induced trajectory approaches the reward-accessible trajectory set defined in Section 3.*

Hypothesis 3 (Reinforcement)**.** *The observer's preference over classes of objects adapts to the reward obtained from them.*

Hypothesis 4 (Weak closure)**.** *For some behaviors, the gain of corrective information about the outcome of the behavior is small relative to the strength of the internally generated reward that sustains it.*

Hypotheses 2 through 4 are developed in Sections 3, 5, and 6 respectively.

3 Resonance, Reward, and Proxy Equivalence

Let $\mathcal{O}$ be a set of external objects, each described by a time-dependent signal $o(t)$, such as the movement of a limb, a flag, or a sound. The percept map of organism i is defined through the coupled dynamics

$$\dot{x} = F(x, 0; \theta_i) + G(x, o(t); \theta_i), \tag{3}$$

where G is the coupling by which the signal perturbs the body. The *induced trajectory* $P_i[o]$ is the solution of (3) started from the rest state x_*. The *motor repertoire* M_i is the set of trajectories that (1) can produce under admissible controls. This definition makes the repertoire depend on θ_i^m, so that an intervention on the body's dynamical parameters changes M_i and generally changes $P_i[o]$ as well.

Two functionals on objects are to be kept apart. The *resonance* of object o for observer i is

$$C_i(o) = \sup_{m \in M_i} \kappa\big(P_i[o], m\big), \tag{4}$$

where κ is a coupling measure between trajectories, for instance a normalized cross-correlation or a phase-locking value. The *value* of the object is

$$V_i(o) = \mathcal{R}_T[P_i[o]]. \tag{5}$$

Resonance is a statement about how well the induced trajectory lies within what the body can itself produce. Value is a statement about how that trajectory is rewarded. The two come apart. An observer can synchronize strongly with a threatening stimulus, an irritating sound, or a painful repetitive movement, and in each case C_i is large while V_i is small or negative. A mild resonance can be neutral. Hypothesis 2 does not say that value increases monotonically with resonance. It says that an object has positive value to the extent that the induced trajectory approaches the reward-accessible trajectory set

$$\Lambda_i(\rho) = \{x(\cdot) \in M_i : \mathcal{R}_T[x] > \rho\},$$

which consists of the repertoire trajectories whose reward exceeds the baseline. This is a set of trajectories and not a region of the instantaneous state space, whose counterpart is the set of states visited by its members. The induced trajectory $P_i[o]$ need not belong to M_i, because external forcing can generate trajectories that the body cannot produce autonomously, so the overlap in Hypothesis 2 is a matter of proximity and not of membership. If $\mathcal{R}_T$ is Lipschitz with constant ℓ_R with respect to the uniform metric $d(x, y) = \sup_{t \le T} \|x(t) - y(t)\|$ on the space $C([0, T]; X)$ of continuous trajectories, which contains both repertoire trajectories and induced trajectories, then

$$V_i(o) \ge \rho - \ell_R \, d\big(P_i[o], \Lambda_i(\rho)\big), \tag{6}$$

because for every $x \in \Lambda_i(\rho)$ we have $\mathcal{R}_T[P_i[o]] \ge \mathcal{R}_T[x] - \ell_R \, d(P_i[o], x) > \rho - \ell_R \, d(P_i[o], x)$, and taking the infimum over x gives (6). An induced trajectory close to the reward-accessible set therefore carries value near or above baseline whether or not it lies in the repertoire. The bound expresses only the continuity of reward near rewarding trajectories. That perceiving an attractive object actually produces trajectories near $\Lambda_i(\rho)$ is the empirical content of Hypothesis 2 and is not implied by the bound. Resonance, which measures closeness to the repertoire as a whole, is neither necessary nor sufficient for this.

Equation (5) defines hedonic value. Two further quantities must be kept distinct from it. The first is a motivational weight $m_i(o)$, the incentive salience attached to the object. The second is an observed choice probability $p_i(o)$, which depends on costs, competing goals, and opportunities as well as on motivation, and which has the schematic form $p_i(o) = \sigma\big(\beta m_i(o) - c_i(o)\big)$ with σ increasing, $\beta > 0$, and c_i a cost term. Hedonic value and motivational weight can dissociate, as the literature cited in Section 2 shows, so that an object can be pursued repeatedly while its hedonic value declines, and an object of high hedonic value can be rarely sought. Repeated attention is therefore not by itself evidence of pleasure, and each intervention of Section 7 must state which of the three quantities it measures.

Definition 2 (Proxy equivalence)**.** For organism i, objects o_1, o_2 are *trajectory-equivalent*, written $o_1 \sim_i^{\mathrm{tr}} o_2$, if $P_i[o_1] = P_i[o_2]$, and *reward-equivalent*, written $o_1 \sim_i^{\mathrm{rw}} o_2$, if $V_i(o_1) = V_i(o_2)$.

Proposition 1 (Nesting and factorization)**.** *Trajectory equivalence refines reward equivalence, so that* $\sim_i^{\mathrm{tr}} \subseteq \sim_i^{\mathrm{rw}}$*. Consequently* $V_i = v_i \circ q_i$*, where* $q_i : \mathcal{O} \to \mathcal{O}/\sim_i^{\mathrm{tr}}$ *is the quotient map and* v_i *is a function on the quotient. The set of objects with value exceeding a threshold is a union of fibers of* q_i*.*

Proof. If $P_i[o_1] = P_i[o_2]$ then $V_i(o_1) = \mathcal{R}_T[P_i[o_1]] = \mathcal{R}_T[P_i[o_2]] = V_i(o_2)$, since V_i is a function of the induced trajectory by (5). This gives the nesting. The factorization is the universal property of the quotient, and the last statement follows because the threshold set $\{o : V_i(o) > c\}$ is the preimage under v_i of an interval and the preimage under q_i of that set is saturated by fibers. □

The content of the proposition is in what it permits. Two objects need not resemble each other in any external feature, or in any feature of biological relevance, in order to be equivalent for an observer. They need only induce the same trajectory, or, under the weaker relation, trajectories of the same reward. The partition of $\mathcal{O}$ into value classes is a property of the observer, determined by θ_i, and it differs from observer to observer. This is the formal sense in which an object is a proxy: what the observer values is the induced internal experience, and the object is one among many ways of producing it.

4 Three Classes of Value Function and an Intervention That Separates Them

To say what the resonance account forbids it must be compared with accounts it is meant to differ from. Two competing classes are defined by the way in which the value of an object depends on the observer.

Definition 3 (Motor-independent classes)**.** Let $\varphi : \mathcal{O} \to \Phi$ be a feature map of the object and let ζ_i collect the parameters of observer i other than θ_i^m. A value function is *motor-independent* if it has the form $V_i(o) = \Psi(\varphi(o), \zeta_i)$ for some Ψ that does not depend on θ_i^m. The *template class* T takes $\varphi = Q$, where $Q(o)$ is a vector of externally specified features of biological relevance, such as indicators of fertility or condition. The *receiver-bias class* S takes $\varphi = S$, where $S(o)$ is the output of a fixed sensory filter, so that value is determined by how strongly the object excites the receiver's sensory system, whatever the biological significance of the object [21, 9]. The *resonance class* R takes value to be given by (5), so that it depends on θ_i^m through M_i and P_i.

The definition is more generous to the competitors than a population-invariant template would be. Neither T nor S requires that all observers share the same preferences. Both permit variation through ζ_i, which can encode developmental history, physiological condition, genetic variation, and learning. What they exclude is dependence on the observer's motor dynamics, and that exclusion is the substantive restriction. It is also what makes the classes distinguishable by intervention, as the following proposition shows.

Proposition 2 (Invariance of motor-independent classes)**.** *Let observers* i *and* i' *satisfy* $\zeta_i = \zeta_{i'}$ *and differ only in* θ^m*. If* V *belongs to* T *or* S *then* $V_i(o) = V_{i'}(o)$ *for every object* o*. In particular,*

an intervention that changes θ^m while holding the object, $Q(o)$, $S(o)$, and ζ fixed leaves value unchanged.

Proof. For $V \in \mathsf{T}$ or $V \in \mathsf{S}$ we have $V_i(o) = \Psi(\varphi(o), \zeta_i) = \Psi(\varphi(o), \zeta_{i'}) = V_{i'}(o)$ because Ψ does not depend on θ^m and the other arguments coincide. □

The proposition is nearly a tautology, and its force lies in what it makes testable. The resonance class does not satisfy the invariance, and the failure can be exhibited in an explicit model. Let the relevant degree of freedom of the observer be an oscillator of phase ϑ with natural angular frequency ω, which is a motor parameter in the sense that it records the rhythm at which the body moves, and let the object be a periodic signal of phase ψ and angular frequency ν, for instance the gait or sway of another body or the oscillation of a flag. Suppose the coupling is the lowest-order symmetric one, so that

$$\dot{\vartheta} = \omega + K \sin(\psi - \vartheta), \qquad \dot{\psi} = \nu, \qquad K > 0. \tag{7}$$

The relative phase $\phi = \psi - \vartheta$ then obeys the Adler equation

$$\dot{\phi} = (\nu - \omega) - K \sin \phi. \tag{8}$$

Write $\Delta = \nu - \omega$. If $|\Delta| < K$, Equation (8) has an exponentially stable fixed point with $\sin \phi^* = \Delta/K$ and $\cos \phi^* > 0$, and the observer is phase-locked to the object. At $|\Delta| = K$ the fixed point is a saddle-node and is nonhyperbolic, so locking there is only marginal, and the boundary is included in the locked band as the limit of the locked states. If $|\Delta| > K$ there is no fixed point and ϕ drifts monotonically with mean rate $\operatorname{sgn}(\Delta)\sqrt{\Delta^2 - K^2}$. The set $\{|\nu - \omega| \le K\}$ is the Arnold tongue of the oscillator.

To connect this to value, suppose that the reward system assigns reward rate λ to the locked state, where $\lambda \in \mathbb{R}$ is a valence parameter belonging to θ^r and may be negative, and a rate that decays with the drift rate $\Omega = \sqrt{\max(\Delta^2 - K^2, 0)}$ otherwise. Concretely let the long-run reward rate be $\bar{V}(\nu; \omega) = \lambda\, s(\Omega)$ with $s(0) = 1$ and s strictly decreasing and vanishing at infinity.

Hypothesis 5 (Rewarded locking)**.** *Phase locking of the observer's oscillator to an external rhythm is rewarded, with the reward law $\lambda s(\Omega)$ above.*

Hypothesis 5 is not a consequence of (7), which establishes only the locking region. It is an independent hypothesis about the relation between motor synchronization and hedonic value, and it is the empirical bridge on which the tuning result depends.

Proposition 3 (Frequency tuning)**.** *Under (7) with reward rate $\bar{V}(\nu; \omega) = \lambda s(\Omega)$, $\lambda > 0$, and s strictly decreasing, the long-run value of an object of frequency ν is a function of $\nu - \omega$ alone. It equals λ on the band $|\nu - \omega| \le K$ and is strictly decreasing in $|\nu - \omega|$ outside the band. Consequently, if $\omega' = \omega + \delta$ then $\bar{V}(\nu; \omega') = \bar{V}(\nu - \delta; \omega)$, so that an intervention changing the motor parameter from ω to ω' translates the preferred band of object frequencies by δ while the object, Q, and S are unchanged. If $|\delta| > K$, the object with $\nu = \omega$ has value λ for the original observer and value $\lambda s(\sqrt{\delta^2 - K^2}) < \lambda$ for the modified observer.*

Proof. By the analysis of (8), the locked state exists exactly when $|\nu - \omega| \le K$, in which case $\Omega = 0$ and $\bar{V} = \lambda s(0) = \lambda$. Outside the band, $\Omega = \sqrt{(\nu-\omega)^2 - K^2}$ is strictly increasing in $|\nu - \omega|$, so $\bar{V} = \lambda s(\Omega)$ is strictly decreasing in $|\nu - \omega|$ because s is strictly decreasing. Since Ω depends on ν and ω only through $\nu - \omega$, so does the value, and replacing ω by $\omega' = \omega + \delta$ gives $\bar{V}(\nu; \omega') = \bar{V}(\nu - \delta; \omega)$. For the last statement take $\nu = \omega$, so that $\Delta = 0$ and the original observer is locked with value λ. For the modified observer $\Delta' = \nu - \omega' = -\delta$ and $|\Delta'| > K$, so $\Omega = \sqrt{\delta^2 - K^2} > 0$ and the value is $\lambda s(\Omega) < \lambda s(0) = \lambda$. □

The proposition does not claim that the Adler equation is a model of the nervous system. It is the simplest mathematical object with the property that matters, which is that the rewarded set of external frequencies is defined relative to an internal parameter, and it makes the contrast with T and S exact. A model in T or S can reproduce the tuning curve for a given observer by choosing Ψ appropriately, but by Proposition 2 it cannot reproduce the translation of the curve under intervention on ω unless ζ_i is allowed to depend on ω. A template model whose parameters track the observer's motor parameters has stopped being motor-independent and has adopted the structure of the resonance class, and the remaining question is whether the feature map Q is then doing any explanatory work. That question is empirical and can be addressed by partial dependence, which is a further reason for keeping the three classes distinct.

5 Reinforcement and the Formation of Preference

Suppose the observer's repertoire of rewarding resonance classes is indexed by $k = 1, \dots, N$, and let $w = (w_1, \dots, w_N)$ be the preference weights, with $w_k \ge 0$ and $\sum_k w_k = 1$, that determine how much of the observer's attention and activity is allocated to class k. A minimal reinforcement rule, in the form familiar from evolutionary game theory [10] and from the broader theory of reinforcement [23], is the replicator equation

$$\dot{w}_k = \eta\, w_k \Big(R_k(w) - \sum_j w_j R_j(w) \Big), \qquad \eta > 0, \tag{9}$$

where $R_k(w)$ is the growth signal that drives allocation to class k when the weights are w. The weights w_k are allocation weights, the share of attention and activity directed to class k, and are therefore tied to choice probability and not directly to hedonic value. If the growth signal is realized hedonic reward, the results below apply as stated, and if it is incentive salience they apply with R_k read as motivational weight, so that a dissociation in which motivational weight is sustained while hedonic value declines yields concentration with declining pleasure. In the remainder of this section R_k is the reward rate obtained from class k. If the rewards R_k are constants with a unique maximum, then every trajectory with positive initial weight on the maximizing class converges to the vertex that concentrates all weight on it. Constant reward therefore predicts winner-take-all concentration, which is one route to the rigid fixation of preference. A weaker assumption yields a different outcome. Suppose that returns diminish with use, because of satiation, habituation, or depletion, in the simplest form

$$R_k(w) = r_k - c_k w_k, \qquad c_k \ge 0. \tag{10}$$

Proposition 4 (Concentration and plurality)**.** *If $c_k = 0$ for all k and $r_1 > r_k$ for $k \neq 1$, then every solution of (9) with $w_1(0) > 0$ converges to the vertex $w = e_1$. If $c_k > 0$ for all k, then every solution with $w(0)$ in the interior of the simplex converges to the unique maximizer w^* of the strictly concave function $\Phi(w) = \sum_k (r_k w_k - \frac{1}{2} c_k w_k^2)$ on the simplex. The maximizer has support $\{k : r_k > \lambda^*\}$ and weights $w_k^* = (r_k - \lambda^*)/c_k$ on that support, where λ^* is determined by $\sum_k w_k^* = 1$, so that preference remains plural whenever more than one class has $r_k > \lambda^*$.*

Proof. Consider first $c_k = 0$. For $k \neq 1$, $\frac{\mathrm{d}}{\mathrm{d}t} \log(w_1/w_k) = \eta(r_1 - r_k) > 0$, so $w_k/w_1 \leq (w_k(0)/w_1(0)) e^{-\eta(r_1 - r_k)t} \to 0$, and hence $w_1 \to 1$.

Now let $c_k > 0$ for all k. Then Φ is a strictly concave quadratic on the simplex with unique maximizer w^*, and $\partial\Phi/\partial w_k = r_k - c_k w_k = R_k(w)$. The first-order conditions for the maximizer state that $R_k(w^*) = \lambda^*$ for k in the support S of w^* and $R_k(w^*) \leq \lambda^*$ otherwise, which gives the stated support and weights. For convergence from an interior initial condition, define

$$D(w) = \sum_{k \in S} w_k^* \log \frac{w_k^*}{w_k}.$$

By Jensen's inequality $D(w) \geq -\log \sum_{k \in S} w_k \geq 0$, with $D(w) = 0$ only at $w = w^*$ (equality in Jensen's inequality forces $\sum_{k \in S} w_k = 1$ and $w_k = w_k^*$ on S), and $D(w) \to \infty$ as any w_k with $k \in S$ tends to zero. Along (9),

$$\dot{D} = -\sum_{k \in S} w_k^* \frac{\dot{w}_k}{w_k} = -\eta \sum_{k \in S} w_k^* (R_k - \bar{R}) = -\eta \, \nabla\Phi(w) \cdot (w^* - w),$$

where $\bar{R} = \sum_j w_j R_j$ and $\sum_{k \in S} w_k^* = 1$ was used. By concavity, $\nabla\Phi(w) \cdot (w^* - w) \geq \Phi(w^*) - \Phi(w)$, which is strictly positive for $w \neq w^*$ because the maximizer is unique. Hence $\dot{D} < 0$ away from w^*. The sublevel sets of D are compact subsets of the closed simplex on which w_k is bounded away from zero for $k \in S$, and they may contain boundary points where coordinates outside S vanish, so LaSalle's invariance principle applied to the replicator vector field, which is polynomial and extends to the closed simplex, gives $w(t) \to w^*$. □

The proposition says that return-diminishing feedback is what makes plurality possible, and it does not say that satiation guarantees it. With $c_k = 0$ the dynamics always select a single class. With $c_k > 0$ the equilibrium is plural exactly when more than one class satisfies $r_k > \lambda^*$, which depends on the gaps between the r_k relative to the c_k, so a single class can still win under satiation if its advantage is large. Where the observer's own activity depletes or habituates the reward, as it does when consumption has consequences in the world, plurality is therefore available, and where the reward is generated by a self-sustaining pattern and nothing in the loop diminishes it, the dynamics select a single class. This is a sufficient condition for fixation and not a description of every case, and it is the bridge to the next section: the absence of outcome-dependent feedback is exactly what the closure parameter of Hypothesis 4 measures.

The same section explains how preferences can be both idiosyncratic and stable without being independently specified. The set of classes available is determined by θ^m through the

induced trajectories of Section 3, the initial weights are determined by history, and the dynamics (9) amplify small initial differences when satiation is weak. A particular movement, sound, texture, or visual pattern can become rewarding through the bodily dynamics it recruits and then become dominant through repetition, even when it bears no relation to any biological function that would be invoked to explain a preference for it.

6 Closure and Sphexishness

Sphexishness is a failure of corrective feedback and should be defined as such, not as preference concentration. A preference can become extremely strong while remaining sensitive to outcomes, and a weak tendency can be rigid if its execution ignores relevant changes in the environment. The relevant structure is the gain with which information about the outcome of an activity enters the dynamics that generate the activity. Let

$$\dot{x} = F(x) + \alpha G(x, o) + \gamma H\big(x, e(t)\big), \tag{11}$$

where $\alpha \geq 0$ is the strength of the resonance coupling to the perceived object, $e(t)$ is information about the outcome in the environment, H is the corrective response to that information, and $\gamma \geq 0$ is the gain with which outcome information enters. The unperturbed system $\dot{x} = F + \alpha G$ is the case $\gamma = 0$, in which the organism sustains its activity entirely from within. That case alone is not informative, since with $\gamma = 0$ the trajectory is trivially independent of $e(t)$. What is needed is a statement about the robustness of an internally sustained rewarding state to a nonzero but bounded γ, because that is what determines whether the persistence regime is an idealization of measure zero or an open region of parameter space.

Assume that the unperturbed system has a compact set Γ that is exponentially stable, and let $W \geq 0$ be a continuously differentiable function with $\Gamma = \{W = 0\}$, together with constants $\kappa > 0$, $w^* > 0$, and $L > 0$ such that on the sublevel set $\Omega^* = \{W \leq w^*\}$ we have

$$\nabla W \cdot \big(F + \alpha G\big) \leq -\kappa W, \qquad \|\nabla W\| \leq L. \tag{12}$$

Converse Lyapunov theorems guarantee the existence of such a W near an exponentially stable equilibrium or set [15]. Let h bound the corrective response, $\|H(x, e)\| \leq h$ for all admissible e and $x \in \Omega^*$, and let r_Γ be the reward rate on Γ, with reward Lipschitz in W, so that $r(x) \geq r_\Gamma - \ell W(x)$ on Ω^*.

Definition 4 (Persistence regime)**.** System (11) is in the *persistence regime* if $\gamma L h < \kappa w^*$.

Proposition 5 (Persistence of the self-sustained attractor)**.** *In the persistence regime, for every admissible outcome signal $e(\cdot)$, every solution starting in Ω^* remains in Ω^* and satisfies $\limsup_{t\to\infty} W(x(t)) \leq \gamma L h/\kappa$. Consequently the long-run average reward satisfies*

$$\liminf_{T\to\infty} \frac{1}{T} \int_0^T r(x(t))\,\mathrm{d}t \;\geq\; r_\Gamma - \ell\,\frac{\gamma L h}{\kappa},$$

whatever outcome signal the environment supplies.

Proof. Along a solution in Ω^*, (11) and (12) give

$$\frac{\mathrm{d}}{\mathrm{d}t}W(x(t)) = \nabla W \cdot (F + \alpha G) + \gamma \nabla W \cdot H \leq -\kappa W + \gamma L h.$$

By the comparison lemma [15],

$$W(x(t)) \leq e^{-\kappa t}W(x(0)) + \frac{\gamma L h}{\kappa}(1 - e^{-\kappa t}) \leq \max\Big\{W(x(0)), \frac{\gamma L h}{\kappa}\Big\}.$$

If $W(x(0)) \leq w^*$ and $\gamma L h/\kappa < w^*$ the right-hand side is at most w^*, so the solution cannot leave Ω^*, which justifies the use of (12) along the whole solution. The lim sup bound follows by letting $t \to \infty$. For the reward bound, $r(x(t)) \geq r_\Gamma - \ell W(x(t))$, and averaging with the lim sup bound on W gives the stated inequality. □

The result is a sufficient condition and establishes persistence of the rewarding state, which is weaker than behavioral rigidity. Remaining near an attractor is not the same as executing the same behavior whatever the outcome. If Γ is a limit cycle, outcome feedback could alter the phase, timing, or sequence of the action while leaving the state close to Γ, and the proposition would still hold. Crossing the threshold $\gamma L h = \kappa w^*$ means only that persistence is no longer guaranteed, and does not establish that behavior is flexible. The persistence regime is an open set in the (α, γ) plane and is not confined to $\gamma = 0$, so rigidity in this sense is a tendency of a region of parameter space and not an idealization, in agreement with the more careful reading of the wasp literature in which rigidity is a matter of degree [14].

Rigidity requires a separate statement about an observable. Let $a : X \to A$ be an action map into a metric space (A, d_A), and for outcome signals e_1, e_2 applied from the same initial state define the outcome sensitivity

$$S_\gamma = \sup_{e_1, e_2} \limsup_{T \to \infty} \frac{1}{T} \int_0^T d_A\big(a(x_{e_1}(t)), a(x_{e_2}(t))\big)\, \mathrm{d}t. \tag{13}$$

Behavior is *ε-rigid* if $S_\gamma \leq \varepsilon$, and a system is *sphexish at tolerance ε* if it is in the persistence regime and is ε-rigid.

Proposition 6 (Bound on outcome sensitivity)**.** *Suppose $C \subseteq \Omega^*$ is convex and forward invariant for (11) under all admissible e, that $\tilde{F} = F + \alpha G$ is contractive on C with rate $\mu > 0$, meaning $\langle x - y, \tilde{F}(x) - \tilde{F}(y)\rangle \leq -\mu\|x - y\|^2$ for $x, y \in C$, that a is Lipschitz with constant ℓ_a, and that $\|H\| \leq h$ on C. Then for any admissible e_1, e_2 and common initial state in C,*

$$\|x_{e_1}(t) - x_{e_2}(t)\| \leq \frac{2\gamma h}{\mu}(1 - e^{-\mu t}) \quad \text{for all } t \geq 0,$$

and consequently $S_\gamma \leq 2\ell_a \gamma h/\mu$.

Proof. Let $\delta = x_{e_1} - x_{e_2}$. Then

$$\tfrac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\delta\|^2 = \langle \delta, \tilde{F}(x_{e_1}) - \tilde{F}(x_{e_2})\rangle + \gamma\langle \delta, H(x_{e_1}, e_1) - H(x_{e_2}, e_2)\rangle \leq -\mu\|\delta\|^2 + 2\gamma h\|\delta\|,$$

so the upper right Dini derivative of $\|\delta\|$ satisfies $D^+\|\delta\| \le -\mu\|\delta\| + 2\gamma h$. With $\delta(0) = 0$ the comparison lemma gives the stated bound. Since $d_A(a(x), a(y)) \le \ell_a \|x - y\|$, the time average in (13) is at most $\ell_a \, 2\gamma h/\mu$. □

Contraction excludes a limit cycle with neutral phase, which is the case in which Proposition 5 holds while action can still differ. For a limit cycle the bound applies only to an action map that does not depend on the phase, or after the phase has been controlled by a separate mechanism. Together the two propositions separate persistence of the rewarding state from insensitivity of the action it generates.

The bound is a statement about differences between trajectories exposed to different outcome signals from the same initial state. It does not by itself show that the behavior is maladaptive, since an action may rightly be invariant to outcome differences that do not matter. Distinguishing rigidity from appropriate invariance requires an independent account of which outcome differences warrant different actions, which can be supplied by an outcome-dependent optimal action $a^\dagger(e) \in A$. Behavior is then *maladaptively rigid at tolerance* ε if $S_\gamma \le \varepsilon$ while $\sup_{e_1,e_2} d_A(a^\dagger(e_1), a^\dagger(e_2)) > \varepsilon$. The mathematics fixes the first condition, and the second is a biological input.

The reinforcement of Section 5 enters through the coupling strength α, and the link can be made explicit. Treat α as a slow variable on $[0, \alpha_{\max}]$ that is raised by reward above baseline and decays at rate $\delta > 0$,

$$\dot\alpha = \varepsilon\big(\bar r(\alpha) - \rho\big) - \delta\alpha, \tag{14}$$

with a projection that keeps α in the interval. Here $\bar r(\alpha)$ is the long-run reward rate on the attractor $\Gamma(\alpha)$ of the fast system, which under a separation of time scales, taken here as an assumption, stands in for the value $V_i(o)$ of the perceived object. Write $q = \varepsilon/\delta$.

Proposition 7 (Reinforcement of coupling)**.** *Suppose $\bar r$ is continuous, nondecreasing, and concave on $[0, \alpha_{\max}]$ with $\bar r(0) > \rho$, so that the state sustained from within already exceeds baseline, as in Hypothesis 1, and suppose $\alpha_{\max} > q(\bar r(\alpha_{\max}) - \rho)$. Then (14) has a unique equilibrium $\alpha^\circ(q) \in (0, \alpha_{\max})$, which is globally attracting on $[0, \alpha_{\max}]$, and α° is nondecreasing in q over the values of q satisfying the endpoint condition. If in addition $\kappa(\alpha)w^*(\alpha)$ is nondecreasing in α, then for fixed γ, L, and h, once the persistence condition $\gamma L h < \kappa(\alpha^\circ(q))w^*(\alpha^\circ(q))$ holds at some q it holds at every larger q.*

Proof. Write $\dot\alpha = -\delta\varphi(\alpha)$ with $\varphi(\alpha) = \alpha - q(\bar r(\alpha) - \rho)$. Since $\bar r$ is concave and $q > 0$, φ is convex. Moreover $\varphi(0) = -q(\bar r(0) - \rho) < 0$ and $\varphi(\alpha_{\max}) > 0$ by the endpoint condition. A zero exists in $(0, \alpha_{\max})$ by the intermediate value theorem, and it is unique: if $\alpha_1 < \alpha_2$ were two zeros there, convexity on $[0, \alpha_2]$ would give $\varphi(\alpha_1) \le (1 - \alpha_1/\alpha_2)\varphi(0) + (\alpha_1/\alpha_2)\varphi(\alpha_2) < 0$, a contradiction. Call the zero α°. Then $\varphi < 0$ on $[0, \alpha^\circ)$, and convexity gives $\varphi > 0$ on $(\alpha^\circ, \alpha_{\max}]$. Hence $\dot\alpha > 0$ for $\alpha < \alpha^\circ$ and $\dot\alpha < 0$ for $\alpha > \alpha^\circ$, which gives global attraction. For monotonicity let $q < q'$ both satisfy the endpoint condition. Since $\alpha^\circ(q) = q(\bar r(\alpha^\circ(q)) - \rho) > 0$ we have $\bar r(\alpha^\circ(q)) > \rho$, so

$$\varphi(\alpha^\circ(q); q') = \alpha^\circ(q) - q'\big(\bar r(\alpha^\circ(q)) - \rho\big) \le \alpha^\circ(q) - q\big(\bar r(\alpha^\circ(q)) - \rho\big) = 0.$$

As $\varphi(\cdot\,; q') \leq 0$ exactly on $[0, \alpha^\circ(q')]$, this gives $\alpha^\circ(q) \leq \alpha^\circ(q')$. The last statement follows because κw^* is nondecreasing in α and α° is nondecreasing in q. □

The proposition supplies the missing connection. Attraction produces reward, reward raises the coupling strength through (14), and the raised coupling widens the persistence regime of Proposition 5, so that the same variable links preference formation to behavioral rigidity. The result remains conditional on the time-scale separation, on the concavity of $\bar{r}$, and on the monotonicity of κw^*, none of which is derived here, and it describes the route from perceptual attraction to rigidity only for behaviors in which the closure of Hypothesis 4 is weak.

Two cautions limit the scope of the claim. The first is that not all stereotyped behavior is reward-driven. Fixed action patterns can persist because of inherited control architecture, environmental elicitation, or learned sequencing without any hedonic reinforcement [24], and the present analysis covers only the subset sustained by self-generated reward. The second is that the term sphexish is borrowed from a story whose empirical basis is contested, and the regime defined above is a mathematical structure, not a claim about the wasp [14, 3].

7 Identifiability and Experimental Discrimination

A theory of this kind risks becoming an alternative vocabulary for familiar observations about learning, pleasure, and individual preference. It is protected against that outcome only if it forbids something. The three classes forbid different things, and Table 1 states the qualitative predictions of the restricted classes defined in Section 4 under four interventions. The entries describe the classes as restricted, not every model that could be built, since a sufficiently flexible model can reproduce almost any single result.

Intervention	Template T	Receiver bias S	Resonance R
Motor training that changes θ^m, stimuli held fixed	No change	No change	Preferred band shifts
Passive sensory exposure with no motor involvement	No change	Change through sensory adaptation	Little or no change
Change in biologically relevant features Q, induced trajectory held fixed	Change	No change	No change
Removal of outcome feedback	No change	No change	Persistence and rigidity increase in the regimes of Section 6

Table 1: Qualitative predictions of the three restricted classes under four interventions.

The table makes the most discriminating test the first row. It requires manipulating the observer's motor dynamics independently of the apparent biological significance of the stimulus, for instance by training a particular rhythmic motor pattern and then measuring preference among movement or sound patterns that differ in frequency but not in any presumed fertility

or condition cue. The prediction of Proposition 3 is a translation of the preferred band by the amount of the shift in the trained frequency. Controls are needed for familiarity, perceptual fluency, novelty, associative conditioning, and general arousal, since each of these can shift preference without involving resonance. Reversible perturbations of the motor system, or controlled training, are cleaner than injury, which alters perception, stress, and reward processing simultaneously. The causal logic follows the standard treatment of interventions in the causal-inference literature [17, 18]. The fourth row holds only under the conditions of Proposition 7, namely the separation of time scales, the concavity of $\bar{r}$, and the monotonicity of κw^*.

The classes are observationally equivalent in an important case. If all observers in a sample share the same motor parameters, then θ^m is constant, every value function is a function of the object alone, and any of the three classes can be fitted to the data. Correlational surveys of what is found attractive in a population of similarly built bodies therefore cannot separate them. Variation in θ^m, whether natural, trained, or perturbed, is required, and the quality of the discrimination is limited by how well θ^m can be measured and manipulated independently of ζ. A further limitation is that the interventions identify the dependence of value on motor parameters but do not by themselves separate hedonic impact from incentive salience, so the reading of r stated in Section 2 must be supported by measures that distinguish the two.

8 Relation to Selection

Nothing in the formalism excludes an evolutionary origin for the components. The reward system, the capacity for motor simulation of perceived movement, and the coupling G between them are all candidates for adaptation, and the study of auditory-motor and sensorimotor coupling in other domains treats them as such [28]. Perception-action coupling in the sense of common coding, in which perceived events and planned actions share a representational domain, supplies a long-standing precedent for the coupling [19, 6, 20, 26, 13], although none of that literature shows that motor resonance as such produces pleasure, and some of it explicitly cautions against treating the neural evidence as conclusive. The point of the present framework is therefore not to deny that selection shaped the machinery but to ask what selection must specify about the objects the machinery responds to.

The three classes answer differently. The template class says that selection specifies the target set through Q. The receiver-bias class says that selection specifies the sensory filter and that the target set is a by-product of its tuning. The resonance class says that selection specifies the reward region, the motor repertoire, and the coupling, and that the target set is then determined by the observer's body and history. The last two are close in spirit, since both deny that targets are individually specified, and the difference lies in what fixes the targets. Receiver bias fixes them in the sensory system and so predicts invariance under motor intervention, as in Proposition 2. Resonance fixes them in the relation between the percept and the body's own dynamics and predicts translation under motor intervention, as in Proposition 3. The claim that resonance contributes beyond motor-independent sensory bias is therefore the claim that the first row of Table 1 comes out as the resonance column predicts, with a qualification stated next.

The qualification concerns the strongest competitor. Sensory and motor systems are reciprocally coupled, and an evolutionary sensory-bias account can incorporate motor-dependent tuning without abandoning its evolutionary explanation. A shift of preferred frequency under motor training would refute the restricted class S defined in Section 4, but it would not refute sensory-bias theory in general. Such a theory shares with the resonance class the dependence of value on θ^m, though not necessarily self-triggerable reward, reward through induced trajectories, or proxy equivalence, since motor dependence can enter through other pathways, and the first row of Table 1 does not by itself separate it from the account developed here. What separates them is the remainder of the conjunction: reward that is reachable from internal states, proxy equivalence defined relative to the induced trajectory, and reinforcement-dependent closure. The distinctive content of the account is that conjunction and not motor dependence alone, and tests of its other components, such as persistence of preference after outcome feedback is removed, are needed in addition to the tuning experiment.

A minimal amount of specified content is compatible with the framework. Early orienting biases, such as a preference for certain patterns in infancy, are specified content that the model can absorb as part of ζ_i or as initial conditions for the weights w without altering the results. The stronger form of the hypothesis is that the specified part is small and that the remainder of the target set arises through drift amplified by reinforcement, as in Section 5. The weaker form, in which the specified part is large and resonance contributes only a fine structure, is not excluded by the formalism and would be found by experiment to be the case if the first-row interventions produced only small shifts.

9 Extension to Rhythm, Sound, and Gait

The Adler model of Section 4 is a model of any case in which an external periodic signal couples to an internal oscillator, and it applies directly to rhythm. A growing body of work establishes that the perception of musical rhythm recruits motor regions of the brain, that neural activity entrains to beat and meter, and that peak emotional responses to music are accompanied by dopamine release in the reward system [28, 4, 22, 12]. These findings do not establish that motor resonance causes pleasure, but they make that causal relation an empirically meaningful hypothesis for one domain in which both the resonance and the reward are measurable. They also supply the strongest present test of Proposition 3 in the auditory case, since the preferred tempo or groove should depend on the listener's own motor tempo in the way the translation property predicts. The mimetic hypothesis in music scholarship, which holds that listening involves covert imitation of the sound-producing actions, is a closely related proposal about the same coupling [5].

Gait is the case that motivated the argument. A walking body produces a periodic signal, the movement of the pelvis and limbs, with a frequency and an amplitude, and an observer has an internal oscillator corresponding to the rhythm of the observer's own locomotion. The framework predicts that attraction to a perceived gait depends on how well that gait lies within the Arnold tongue of the observer's own locomotor oscillator, that the dependence is graded and not categorical, and that the preferred band moves with training of the observer's own gait.

It also predicts that the same coupling attaches to objects with no biological significance, such as a flag or flame whose oscillation falls within the band. The prediction concerns frequency and phase relations and not any claim about specific human data, none of which is asserted here.

10 Limitations and Conclusion

The formalism leaves open several matters that determine whether the hypotheses are true. The form of the percept map G, which is the part of the model that would need to be identified from neural and behavioral data, is not specified beyond the Adler example. The measure κ of resonance and the reward functional r must be estimated, and the second requires a decision about which component of reward is at issue. The Adler model is a one-dimensional caricature of phase coupling, and richer repertoires would involve many interacting oscillators and a more complicated analysis of locking. The persistence and rigidity bounds are sufficient conditions derived under the assumptions of an exponentially stable attractor and, for rigidity, contraction, and real systems may be rigid or flexible for reasons the analysis does not cover. The monotonicity assumption linking reinforcement to the width of the regime is an assumption and has not been derived. None of these limitations undermines the main structural claim, which is that the three classes make different predictions under intervention, but together they mean that the formal results are a framework for experiments and not a substitute for them.

The paper has argued that a self-triggerable reward system supports a class of account in which the objects of attraction need not be individually specified by selection. Under that class, an object is valued through the trajectory it induces in the observer, objects are proxies whose equivalence is relative to the observer, preference concentrates or remains plural depending on satiation, and a self-sustained rewarding attractor persists, and under contraction generates action that is insensitive to outcome feedback, in an identifiable region of parameter space. The same internal mechanism therefore supports flexible attraction, idiosyncratic preference formation, and rigid fixation, depending on how its dynamics are coupled to external feedback. The most important result to pursue empirically is not that arbitrary objects can become attractive, which the receiver-bias literature already allows, but that translation of the preferred band under intervention on the observer's own motor dynamics distinguishes resonance from the motor-independent classes. If that effect is absent, given successful manipulation and measurement of the motor parameter, the Adler-based tuning model is refuted in the domains tested, although the resonance class more generally is not, since some motor interventions may leave the reward-relevant induced trajectories unchanged. If it is present, it refutes the motor-independent classes as defined here but not sensory-bias theory in general, which can accommodate motor-dependent tuning, and what then distinguishes the present account is the conjunction of self-triggerable reward, induced sensorimotor trajectories, reward-relative proxy equivalence, and reinforcement-dependent closure. Where that conjunction holds, the question of what evolution must specify about the objects of attraction has a more restricted answer than the standard accounts assume.

References

[1] Berridge, K. C. (2007). The Debate over Dopamine's Role in Reward: The Case for Incentive Salience. *Psychopharmacology*, 191, 391–431.

[2] Berridge, K. C., and Robinson, T. E. (1998). What Is the Role of Dopamine in Reward: Hedonic Impact, Reward Learning, or Incentive Salience? *Brain Research Reviews*, 28(3), 309–369.

[3] Brockmann, H. J. (1985). Provisioning Behavior of the Great Golden Digger Wasp, *Sphex ichneumoneus* (L.) (Sphecidae). *Journal of the Kansas Entomological Society*, 58(4), 631–655.

[4] Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008). Listening to Musical Rhythms Recruits Motor Regions of the Brain. *Cerebral Cortex*, 18(12), 2844–2854.

[5] Cox, A. (2016). *Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking.* Indiana University Press.

[6] Decety, J., and Grèzes, J. (1999). Neural Mechanisms Subserving the Perception of Human Actions. *Trends in Cognitive Sciences*, 3(5), 172–178.

[7] Dennett, D. C. (1984). *Elbow Room: The Varieties of Free Will Worth Wanting.* MIT Press.

[8] Dixson, A. F. (1998). *Primate Sexuality: Comparative Studies of the Prosimians, Monkeys, Apes, and Human Beings.* Oxford University Press.

[9] Endler, J. A., and Basolo, A. L. (1998). Sensory Ecology, Receiver Biases and Sexual Selection. *Trends in Ecology and Evolution*, 13(10), 415–420.

[10] Hofbauer, J., and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics.* Cambridge University Press.

[11] Hofstadter, D. R. (1985). *Metamagical Themas: Questing for the Essence of Mind and Pattern.* Basic Books.

[12] Janata, P., Tomic, S. T., and Haberman, J. M. (2012). Sensorimotor Coupling in Music and the Psychology of the Groove. *Journal of Experimental Psychology: General*, 141(1), 54–75.

[13] Jeannerod, M. (2001). Neural Simulation of Action: A Unifying Mechanism for Motor Cognition. *NeuroImage*, 14(1), S103–S109.

[14] Keijzer, F. (2013). The Sphex Story: How the Cognitive Sciences Kept Repeating an Old and Questionable Anecdote. *Philosophical Psychology*, 26(4), 502–519.

[15] Khalil, H. K. (2002). *Nonlinear Systems* (3rd ed.). Prentice Hall.

[16] Olds, J., and Milner, P. (1954). Positive Reinforcement Produced by Electrical Stimulation of Septal Area and Other Regions of Rat Brain. *Journal of Comparative and Physiological Psychology*, 47(6), 419–427.

[17] Pearl, J. (2009). *Causality: Models, Reasoning, and Inference* (2nd ed.). Cambridge University Press.

[18] Peters, J., Janzing, D., and Schölkopf, B. (2017). *Elements of Causal Inference: Foundations and Learning Algorithms*. MIT Press.

[19] Prinz, W. (1997). Perception and Action Planning. *European Journal of Cognitive Psychology*, 9(2), 129–154.

[20] Rizzolatti, G., and Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, 27, 169–192.

[21] Ryan, M. J. (1998). Sexual Selection, Receiver Biases, and the Evolution of Sex Differences. *Science*, 281(5385), 1999–2003.

[22] Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., and Zatorre, R. J. (2011). Anatomically Distinct Dopamine Release during Anticipation and Experience of Peak Emotion to Music. *Nature Neuroscience*, 14, 257–262.

[23] Sutton, R. S., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.

[24] Tinbergen, N. (1951). *The Study of Instinct*. Oxford University Press.

[25] Tomasello, M. (2022). *The Evolution of Agency: Behavioral Organization from Lizards to Humans*. MIT Press.

[26] Wilson, M., and Knoblich, G. (2005). The Case for Motor Involvement in Perceiving Conspecifics. *Psychological Bulletin*, 131(3), 460–473.

[27] Wooldridge, D. E. (1963). *The Machinery of the Brain*. McGraw-Hill.

[28] Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the Brain Plays Music: Auditory–Motor Interactions in Music Perception and Production. *Nature Reviews Neuroscience*, 8, 547–558.