

# Not so Basic Instinct: Learning, Evolution, and the Behavioural Hypercycle

Douglas Roy

Australia, UNSW; douglas.roy@student.unsw.edu.au

## Abstract

Traditional ideas about cognitive evolution often treat learning and instinct as zero-sum substitutes, where increased plasticity necessarily displaces hard-wired routines. We challenge this "substitution-only" view using economic concepts and by considering evolutionary implications of "behavioural hypercycles": i.e., teams of modular component behaviours coordinated toward a functional task. This shows that, while intermediate brain sizes can indeed favour the substitution of instinct for flexible learning, larger nervous systems trigger a sort of "Income Effect" which changes the optimal allocation between how growing neural resources are dedicated between learning and instinct. Rather than displacing one another, sophisticated learning and extensive instinctive repertoires can evolve as adaptive complements under identifiable conditions. We further show that this trade-off is likely level-specific: evolution preferentially canalizes reusable, high-burden primitives (i.e., genetically assimilating behavioural components that are especially costly or difficult to learn) while leaving task-specific links plastic (e.g., subject to goal-directed control). Our analysis suggests that hypercyclic organization is a fundamental principle of complex agency, where instinct provides the reliable scaffold that makes sophisticated learning affordable. Our model is consistent with several lines of evidence (behavioural, genetical, neurological), likely applies broadly to any animals capable of complex behaviour, and points to a range of empirically testable predictions.

**Keywords:** instinct; ethology; adaptation; hypercycle; baldwin effect; instrumental learning; pavlovian learning; behavioural sequences; exaptation; rotten kid theorem; habits; subsumption architecture

---

## 1. Introduction

The smartest moves are not always made by the cleverest creatures. Aesop's fable of the fox and the hedgehog juxtaposes two modes of problem-solving: one imaginative and wide-ranging, the other single-minded and instinctive. As Stephen Jay Gould (2003) put it, these represent alternative tendencies to "diversify and colour" versus "intensify and cover." In nature, such cognitive styles may be seen as spanning a continuum between the flexible and reflective at one extreme and the fixed and rigid on the other (Dennett, 1984). Both can produce strikingly sophisticated behaviour. Caddisfly larvae, for instance, construct underwater shelters by binding grains of sand, plant fragments, or shell pieces with silk that hardens on contact with water. These larvae sort and arrange materials selectively according to size and weight, using inherited motor routines refined by evolution rather than organized by imaginative planning. Interestingly, as Dawkins (1982) observed, if we saw a more intelligent animal such as a dolphin or chimpanzee performing such underwater engineering, we would marvel at its ingenuity. Yet the caddisfly achieves all this and more with a nervous system whose total weight is a fraction of a milligram. Something similar can be said for orb-weaving spiders, which execute hundreds of geometrically precise movements to spin complex webs, or for mason bees, potter wasps, and termites whose nests similarly embody intricate architectural acumen without insight (Gould & Gould, 2007). These examples show that behavioural complexity need not imply intelligence: remarkable organization can emerge from instinctive programs evolved (somehow) by blind selection acting on genetic variation.

Dawkins went on to suppose that this means we should perhaps be *more* impressed by the caddisfly than by any dolphin making such moves, because the caddisfly's performance must be explained by evolution alone. An implication is that there is a greater burden on adaptationist reasoning to explain such instinctive feats than the behaviour of a generally intelligent organism. For the dolphin, nearly any specific act of problem-solving may be explicable under its general cognitive capacities. Once the deferred question as to how the dolphin evolved to be so clever in the first place is answered, that explanation can subsume any specific expressions of this cleverness. But for the caddisfly, no such shortcut exists. The aims of this paper are to consider how complex and adaptive behaviours may have evolved in animals like the caddisfly and what this mechanism suggests about the relationship between instinct, learning, and intelligence across taxa more generally.

### 1.1. *The Baldwin Effect*

An idea relevant to how learning can influence the evolution of instinct is the Baldwin effect. To illustrate, imagine some environmental challenge, such as the presence of a nutritious nut or shellfish being abundant but difficult for other animals to handle effectively for consumption (i.e., crack open its hard external shell to expose the edible insides). Some individuals in this environment may learn a "Good Trick": a behaviour such as by dropping the shell from a great height or placing it on a flat rock and hammering it with a sharp stone (Dennett, 1995). Suppose that consuming the nuts yields a high fitness benefit to any animal learning the behaviour, and that the habitat is sufficiently stable so that it continues to feature as an important food source over many generations. If there is genetic variation in acquiring this technique, selection may then favour genetic variants biasing individuals towards learning it more rapidly, such as by selection favouring alleles that happen to code the brains of individuals in ways that (somehow) make them more likely to pick up and move the nut, more likely to play with stones, etc., thus making rediscovery of the Good Trick that bit more probable over successive generations. By favouring such variants as they arise, selection can progressively reduce the environmental input required to acquire the behaviour (genetic assimilation), eventually rendering it instinctive.

An interesting implication of this is that traits may evolve more rapidly than would otherwise occur without learning (Dennett, 1995; Maynard Smith, 1987). This is because the probability of an individual initially being born with precisely the necessary genetic configuration to innately know the Good Trick is vastly different from the likelihood that there may be additive genetic variation influencing how readily the behaviour is learned. Yet the former is what would be required if the capacity for learning did not exist. In the selective landscape metaphor, a genotype completely encoding the Good Trick is like an isolated telegraph pole that selection cannot climb, whereas learning can enable adaptation by exposing to selection infinitesimal approximations of that "ideal" genotype, enabling the adaptation to evolve it more like climbing a gentle hill towards its summit (Dennett, 1995) with modelling confirming that even basic reinforcement learning can greatly accelerate subsequent adaptation (Hinton & Nowlan, 1987).

There is of course an alternative path selection might take to genetic assimilation. It may simply favour enhancing learning capacity itself. That is, rather than encode this specific Good Trick (intensify and cover), the species may evolve a progressively greater capacity for generally learning Good Tricks (diversify and colour). As thus told, it is tempting to suppose that this presents a sort of fork in the road, where some lineages evolve down a path towards robotic repertoires while others towards flexible and greater learning capabilities. Which path is taken has been suggested to hinge on the degree to which characteristics of the environmental challenge remain consistent across generations, with highly stable environments favouring genetic assimilation and fluctuations across generations and throughout a species range promoting plasticity (Stephens, 1991). We demonstrate below that these paths are not as mutually exclusive as they might appear but may in fact become intertwined and converge in high-capacity systems.

### 1.2. *The Structure of Behaviour*

To understand this, we must examine the structure of the behaviours in question. Many complex instincts and learned behaviours share key structural properties. Namely, instincts are often characterized by the following common motif: a sequence of functionally linked behaviours that unfold in a fixed order once triggered by the detection of a specific stimulus (Thorpe, 1956). A wasp preparing a nest may perform an invariant pattern of digging, provisioning, and sealing. A spider laying out its web follows a highly stereotyped trajectory, modulated only by spatial cues. A greylag goose that begins to retrieve a displaced egg continues the motion to completion even if the egg is removed mid-sequence. Each case shows an apparently hard-wired algorithm for solving a recurring ecological problem. The interesting thing about this structure is that it mirrors patterns long familiar to animal trainers and psychologists, where sequences of several distinct actions can be “chained” together through conditioned reinforcement. Consequences of the completion of each learned response or behavioural step becomes the discriminative cue for the next in the series, with the initial link in the chain being prompted by an environmental stimulus and the terminal behaviour typically being followed by a food reward or some other external source of reinforcement (Skinner, 1938). Textbook examples of chains can include a pigeon pecking, turning, and pulling a string to receive grain, or a rat pressing a lever and then climbing a ladder to reach food, or pulling a chain to then access a lever which can then be pressed to earn a reward.

Denoting each distinct behavioural component in such a sequence as  $R_i$ , we can compare instinct to a sequential behaviour chain that can be entirely learned, that is, conditioned by “chaining” procedures (Hull, 1942).

Sign Stimulus  $\rightarrow R_1 \rightarrow R_2 \rightarrow R_3 \rightarrow R_4 \rightarrow R_5 \rightarrow$  Consummatory Act/Stimulus

$S_1 \rightarrow R_1 \rightarrow S_2 \rightarrow R_2 \rightarrow S_3 \rightarrow R_3 \rightarrow S_4 \rightarrow R_4 \rightarrow S_5 \rightarrow R_5 \rightarrow O_5$

Where  $S_i$  represents a stimulus that  $R_{i+1}$  is conditioned to be triggered by, which itself is either contingent on execution of  $R_{i-1}$  (except  $S_1$  which is usually controlled by the environment<sup>1</sup>), and  $O_5$  represents the reinforcement delivered upon completion of the chain (usually food). The most commonly used method of chaining is “backward chaining”. This means initially training the animal to perform, say  $S_5 \rightarrow R_5 \rightarrow O_5$  only (e.g., Jerome *et al.*, 2007). When this is fully learned, descriptively what also occurs is that  $S_5$  becomes a conditioned or learned source of reinforcement by association with  $O_5$ . This means perception of  $S_5$  itself becomes capable of reinforcing other associations, at which point the animal is then trained  $S_4 \rightarrow R_4 \rightarrow S_5 \rightarrow R_5 \rightarrow O_5$ , which is then repeated (and so on, and so forth).

The resemblance raises the question as to whether the two forms share a common underlying logic or evolutionary pathway via the Baldwin effect. If a species has already evolved various, individually simple behaviours, complex adaptive behaviour can conceivably arise from the linking of simpler units into ordered sequences if the combination is successful in unlocking greater or new sources of reinforcement. As with learned chains guided by experimenters in the lab, reinforcement could link behaviours together in nature to form functional sequences. Assuming that reinforcement is sufficiently aligned with positive effects on Darwinian fitness, then, when learning repeatedly produces such a behaviour combination, genetic variants supporting its efficient execution may be favoured, gradually embedding once-plastic routines into inherited architecture. In other words, we may consider if complex instincts, such as any conforming to the “fixed action” motif mentioned above, have evolved via the Baldwin effect. Crucially, this could occur in species that need not be capable of generating similar behaviours based on insight or goal-directed processes. To evolve an instinct via chaining and genetic assimilation requires that the ancestors (1) had basic associative conditioning capacities (e.g., stimulus-response learning); (2) were already capable of individually much simpler behaviours; and (3) would at least occasionally execute such behaviours in novel or

---

<sup>1</sup> However, note that some sorts of stimuli important for this sort of learning might even be internal feedback signals produced by the very performance of that behaviour--the sort of “fractional anticipatory goal-responses” of Hullian terminology (Amsel, 1994). For instance, even perception of one’s skeletal responses can feedback as stimuli into conditioned associations (McNaughton, 1989).

challenging situations where serendipitous success would unlock strong sources of reinforcement (these sources presumed here to correspond with positive outcomes on Darwinian fitness).

### 1.3. Aims and Premises

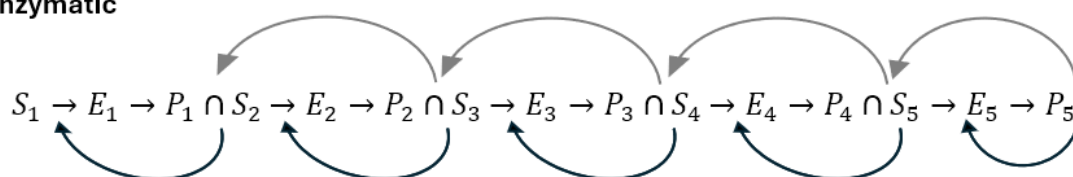
This sort of “complementarity” between simpler behavioural components has been described as a “behavioural hypercycle”. The term “hypercycle” was originally introduced by Eigen and Schuster (1977; 1979) to describe self-reinforcing catalytic networks among molecular replicators. Roy (2017), extended this logic to behavioural and cultural evolution, arguing that behaviours can similarly form and stabilize as autocatalytic systems. This idea can be illustrated by rewriting the above expressions of behaviour to compare with the classic enzymatic hypercycle as in Figure 1<sup>2</sup>. This conceptually shows how learned behavioural chains can, through iterative feedback between performance, learning, and selection, evolve into instinctive sequences. If

(1) behaviours, whether separately evolved or acquired, can be combined by learning into higher order combinations to unlock new benefits, and

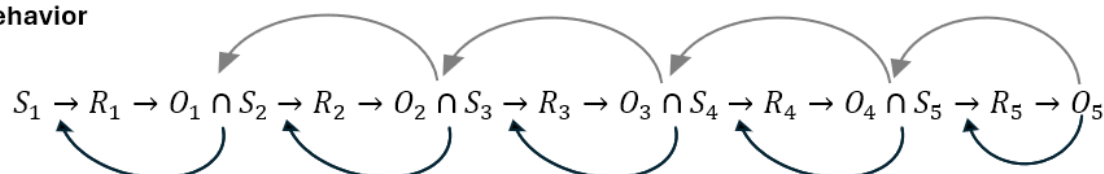
(2) the challenge is recurring, then evolution may lead to the genetic assimilation of the combination,

then such arrangements mean that the total amount of information embodied by the hypercycle and replicated across generations may thereby exceed limits on how complicated a behaviour can otherwise evolve (for further description of this “error catastrophe” problem in the case of behaviour see Appendix A).

#### (a) Enzymatic



#### (b) Behavior



**Figure 1. Diagrammatic representation of hypercycles.** (a) shows a simple hypercycle for five autocatalytic enzyme reactions. The products of each reaction include consequences that increase their own replication. Products also include consequences that enhance the replication of another autocatalytic reaction in the chain (represented by dashed causal arrows). The loop is closed because the products of the fifth reaction feedback to enhance the replication of the first. (b) shows a *behavioural* hypercycle consisting of five stimulus-responses. The products of the fifth behaviour include primary reinforcement that acts to make products of the fourth reaction become secondary reinforcers. Secondary reinforcement thus cascades backwards through the chain, eventually promoting the replication of the first stimulus-response. This iterative reinforcement thus closes the loop (Modified from Figure 3 in Roy, 2017).

<sup>2</sup> We have used the  $\cap$  symbol here because it captures the requirement that some aspects of the effects caused by each behaviour in the chain must overlap with the perceptual stimuli necessary to stimulate the next.

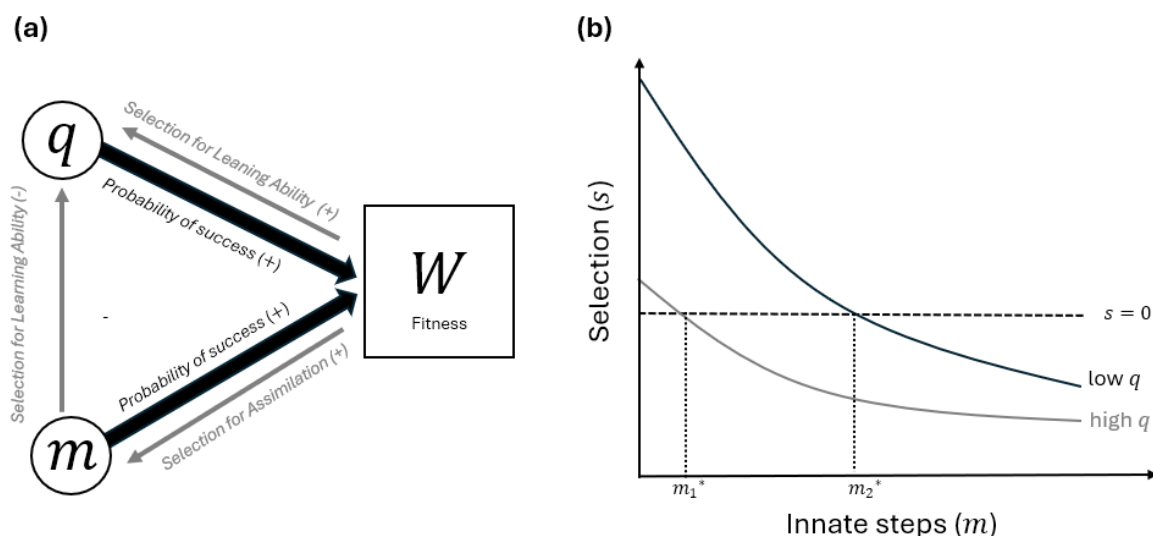
What makes this hypercyclic organisation special? By definition, each component in a successful and stable hypercycle *must* be a specialist. For reasons eminently explored by economic theory, greater efficiencies may be discovered through dividing up a task into specialists (E.g., Roy 1951). This means that teaming up in this way is at least theoretically likelier to produce more reward (higher fitness payoffs) than the sum of its parts otherwise deployed. This principle of specialization and division of labour naturally leads to each component further adjusting, whether by practice effects, coevolving, or both, to achieve still higher efficiencies. The same process would be further augmented by duplication and subsequent further specialization of behavioural components (although potentially counteracted by “hold up” or “coordination costs”).

When put in these terms, instinct and learning can be expected to have complementary (mutually reinforcing) roles in the evolution of one another, with learning ability facilitating the evolution of more instincts by discovery and subsequent genetic assimilation, which in turn may maintain selection for learning ability. This departs from the traditional view of alternative paths, such as evolving to become better at learning instead of relying on instinct, which treat learning and instinct as “substitutes”. This means that it is by no means clear that an advanced capacity for flexible behaviour and learning should be negatively correlated with the number of instincts in an animal’s repertoire and raises the question of whether they might even positively covary in important ways. To help clarify the conditions under which instincts emerge and derive implications for understanding learning and behaviour, the model developed in the next section translates this qualitative idea into a set of dynamical relationships between behavioural performance and evolutionary change.

## 2. The Model

We develop a minimal population-genetic model of a multi-step behavioural hypercycle to show how the Baldwin effect can drive genetic assimilation. We consider behaviour as composed of discrete elements or acts that can be linked into chains. Within an individual’s lifetime, reinforcement strengthens associations between acts that lead to reward, thereby increasing the likelihood that successful sequences will be repeated. Across generations, individuals differ genetically in their predispositions to perform or to learn particular elements of these sequences. When learning yields fitness-enhancing outcomes, selection favours alleles that increase the probability or ease of producing those successful sequences, effectively “hardwiring” portions of what was once learned.

The model thus tracks three coupled processes: (1) behavioural performance, determining the immediate success of sequences; (2) associative learning, which modifies the transition probabilities between behavioural elements; and (3) selection, which alters the population distribution of genotypes that bias learning or performance. Together, these processes form a feedback system akin to the behavioural hypercycle depicted in Figure 1, in which learning catalyses the evolution of instincts and inherited predispositions, in turn, guide the formation of new learned behaviours, as illustrated in Figure 2a.



**Figure 2. Feedback between learning and genetic assimilation in a behavioural hypercycle.** (a) Conceptual flow diagram of the Baldwin-effect dynamics. Learning ability ( $q$ ) and innate encoding ( $m$ ) increase the probability of completing the behavioural sequence and thus fitness (black arrows denote within-lifetime effects). Fitness feeds back through natural selection to alter both traits across generations (grey arrows). Because additional innate steps reduce the marginal value of learning, selection on  $q$  declines as  $m$  increases, creating negative evolutionary feedback between learning and instinct. (b) Selection coefficient  $s$  for an allele that adds one additional innate step as a function of the number of already innate steps ( $m$ ). The horizontal line marks  $s = 0$ . Regions above the line favor assimilation, whereas regions below favor learning. Curves illustrate different learning abilities ( $q$ ); their intersections with  $s = 0$  (labelled  $m^*$ ) indicate candidate evolutionary equilibria.

To keep the model simple, think of the hypercycle as a task made up of an ordered sequence of  $N$  steps. Completing all steps successfully yields a fitness bonus  $B$  — such as better access to food or mates. For clarity, baseline fitness is normalized to 1 and left implicit, meaning that all fitness expressions are measured relative to this baseline. Each step can be either innately encoded or learned during life. We assume one genetic locus per step, with two possible alleles: (1) an *innate* allele, which hardwires the step, and (2), a non-innate allele, which leaves the step to be learned during an individual's lifetime. For tractability, we assume that the probability of learning each step is independent of the others. This is a conservative assumption, since we suspect that relaxing it would strengthen the complementarity between steps in the hypercycle and hence increase assimilation pressure.

In this and subsequent sections, we let  $q$  represent learning ability, which we model as the probability of successfully learning any given non-innate component of the hypercycle. We assume  $q$  (a polygenic or developmental trait) adapts quasi-continuously between rare fixation events of innate steps. All learned steps share this learning probability. Relying on learning comes with two kinds of expenses. There is the per-step learning cost  $\ell(q)$ , which is independent of outcome and rises with  $q$ , reflecting energy, risks, time and other resources spent on each learning attempt. The other type of cost comes from physiological maintenance  $M(q)$ , such as building brain tissue, which grows convexly with increases in  $q$ , representing the expense of maintaining neural or cognitive machinery to support learning<sup>3</sup>.

<sup>3</sup> For tractability we assume that learning and encoding costs are additive across steps and that learning success probabilities are independent. This deliberately removes interference among components. Such effects would primarily rescale marginal costs or benefits and therefore shift quantitative thresholds without altering

Obviously, if  $m$  of the total  $N$  steps are innate, all the remaining  $(N - m)$  steps must be learned to complete the chain. Thus, the total learning cost is

$$\text{learning cost} = \ell(q)(N - m) + M(q).$$

To the extent that behavioural steps do not require learning, they avoid this cost. But we will assume even fully innate steps also incur costs of their own, such as a genetic/developmental cost  $g$  per innate allele, such that the total cost of an instinctive behaviour is

$$\text{Instinct cost} = gm.$$

An individual succeeds only if every step is either innately encoded or successfully learned. With  $m$  innate steps, the probability of producing the complete hypercycle is

$$P(\text{complete}) = q^{N-m}$$

The expected fitness of an individual with learning ability  $q$  and  $m$  innate steps is the baseline fitness for any ordinary individual in the population plus the benefit of successfully “cracking the nut” (i.e., executing the full chain or Good Trick correctly) minus the three types of costs of knowing the Good Trick (learning, neural, and genetic costs), is measured as

$$W(q, m) = 1 + B q^{N-m} - \ell(q)(N - m) - M(q) - gm.$$

To see how selection acts on learning ability, we take the partial derivative of fitness with respect to  $q$  while holding  $m$  constant:

$$\frac{\partial W}{\partial q} = B(N - m) q^{N-m-1} - (N - m)\ell'(q) - M'(q).$$

This captures the balance between the marginal benefit and marginal costs of improved learning ability on fitness. The benefit term is the marginal improvement in the probability of succeeding in producing the hypercycle, scaled by  $B(N - m)q^{N-m-1}$ . Costs marginally increase in per-step learning  $(N - m)\ell'(q)$  and physiological costs  $M'(q)$ .

What level of learning ability should then be adaptive? We treat learning ability as adjusting to its local fitness optimum between rare assimilation events (a standard separation-of-timescales assumption). A singular point  $q^*$  then satisfies

$$B(N - m) q^{N-m-1} = (N - m)\ell'(q^*) + M'(q^*).$$

As  $m$  increases (more steps become innate), both  $N - m$  and  $q^{N-m-1}$  shrink. Hence optimal learning ability  $q^*$  declines as more steps become innate. Intuitively, there is less for learning to do once many steps become genetically specified.

Now consider genetic assimilation. Suppose the resident population has  $m$  innate steps and learning ability  $q$ . A rare mutant arises that makes one additional step innate, raising the number of innate steps to  $m + 1$ . Learning ability is assumed not to change initially. The resident fitness is

$$W_{\text{res}} = 1 + B q^{N-m} - \ell(q)(N - m) - M(q) - gm,$$

while the mutant fitness is:

$$W_{\text{mut}} = 1 + B q^{N-(m+1)} - \ell(q)(N - (m + 1)) - M(q) - g(m + 1).$$

The selection coefficient is the difference:

$$s = W_{\text{mut}} - W_{\text{res}} \\ s = B(q^{N-(m+1)} - q^{N-m}) + \ell(q) - g$$

Factor the difference:

$$q^{N-(m+1)} - q^{N-m} = q^{N-m-1}(1 - q)$$

Thus, the selection coefficient simplifies to

$$s = B q^{N-m-1}(1 - q) + \ell(q) - g.$$

This result isolates three forces: (1) the reliability gain (the benefit from the least reliably learned steps becoming innate); (2) the saved learning cost; and (3) genetic cost. As illustrated in Figure 2b, a new innate allele spreads when

$$B q^{N-m-1}(1 - q) + \ell(q) > g.$$

---

the qualitative regimes identified below. If anything, realistic interference in learning would further favour genetic assimilation, making our results conservative.

The first term is the marginal benefit of improving the success probability ( $B q^{N-m-1}(1-q)$ ). This is larger when  $q$  is low (learning is difficult), and when many steps are already innate (so the residual learned chain is shorter and  $q^{N-m-1}$  is larger). The remaining terms are marginal increases in per-step and maintenance costs:  $\ell(q)$  represents per-step learning cost saved and  $g$  is the genetic cost of encoding a step innately that the marginal benefits from genetically assimilating the step must overcome. As  $m$  increases, the exponent  $N-m-1$  decreases: the marginal probability gain rises (because  $q^{N-m-1}$  increases as the residual chain shortens). Because optimal learning ability  $q^*$  also falls with rising  $m$ , learning becomes less effective. This can boost the gain from assimilation.

Because both the marginal value of learning and the payoff to assimilation depend on how many steps are already innate, feedback may generate non-linear dynamics and multiple locally stable equilibria: that is, these interacting forces create threshold effects. Depending on costs and benefits, the system can settle in one of three plausible regimes. One is a learner-dominated equilibrium, characterised by high learning ability  $q^*$  and where few, if any steps become innate. This occurs when learning is cheap and innate encoding is costly. Another is an assimilation-dominated equilibrium, with low learning ability maintained while many (or all) steps become instinctive. This occurs when learning is expensive, slow, or physiologically burdensome. Finally, there is an intermediate equilibrium, with partial assimilation and moderate learning ability. This occurs when neither purely learned nor purely innate strategies are globally optimal.

Because the marginal payoff to adding an innate step depends nonlinearly on how many steps are already innate, and on how this feeds back into optimal learning ability, the evolutionary trajectory is typically path dependent. Early fixation (or loss) of one or two steps can steer the system toward vastly different equilibria: once a species starts down the path of "intensifying and covering" (instinct), for instance, the marginal value of "diversifying and colouring" (learning) drops so sharply that they may be unlikely to pivot back.

For concreteness and to make these Evolutionarily Stable scenarios explicit, we adopt the simple cost functions  $\ell(q) = aq$  and  $M(q) = bq^2$  (with  $a, b > 0$ ). The stationarity condition  $\frac{\partial W}{\partial q} = 0$  then reduces to

$$B(N-m)q^{N-m-1} = (N-m)a + 2bq,$$

which, after multiplication by  $q$  is

$$B(N-m)q^{N-m} - (N-m)aq - 2bq^2 = 0.$$

This is a polynomial, and  $q^*$  is the root in  $[0, 1]$ . Closed-form solutions exist for small cases (e.g.,  $N-m=1$  gives  $q^* = \frac{B-a}{2b}$ ; if  $N-m=2$  gives  $q^* = \frac{a}{B-b}$ , subject to feasibility), while higher-order cases require numerical solutions. This explicit form makes transparent how neural investment costs ( $a, b$ ), task reward  $B$ , and remaining learned steps  $N-m$  determine the ESS learning ability and thus the balance between learning and genetic assimilation. For small residual chain-lengths  $k = N-m$  this gives a clear intuition: when only one step remains,  $q^*$  increases linearly with  $B$  and is bounded by cost parameter. When two steps remain, a simple rational expression arises. For larger  $k$ , the marginal benefit term  $Bkq^{k-1}$  is highly nonlinear, making the ESS sensitive to  $k$  (i.e., how many steps remain requiring learning). This captures the basic insight that, as more steps become innate, selection for learning ability shifts in important ways. This helps us handle how to expect evolution to proceed once creatures discover new behavioural hypercycles. Parameters  $a, b, N, m$  shape  $q^*$  and so may be useful for predicting when the Good Trick should become instinctive and when learning capacity should remain more crucial.

This model shows how learning ability and instinct interact in a feedback loop shaped by costs, benefits, and task complexity. When learning is cheap and effective, evolution favours flexible learners; when learning is costly or unreliable, genetic assimilation dominates. Intermediate cases produce mixed strategies. Crucially, the trajectory is path-dependent: early changes in the balance between innate and learned steps can lock populations into very different evolutionary outcomes. By formalizing these dynamics, we gain a clearer picture of when behavioural hypercycles remain plastic and when they crystallize into instinct.

### 2.1. Heterogeneous Steps and Greedy Assimilation

We now extend the model by introducing heterogeneity among steps and exploring how these shape the evolutionary dynamics of their ordering. In Eigen's original hypercycle concept, each quasi-species produces a unique product that catalyses the replication of the next quasi-species in a strict sequence. If all quasi-species were identical and interchangeable, the system would collapse into a single autocatalytic entity, losing the mutually reinforcing complementarity that makes hypercycles stable. The advantage of a hypercycle lies in each component playing a distinct, indispensable role in a precise order.

Behavioural hypercycles work similarly: they require a sequence of heterogeneous steps completed to earn a fitness benefit. The current set of innate steps  $m$  has size  $|m|$ , with the remaining steps learned anew each lifetime. For each step  $i$ , learning success probability when non-innate is  $q_i$  (equal or variable across steps). The learning cost saved when a step is innate is  $c_i$ , which may depend on investment in the coevolutionary model but here is treated as given<sup>4</sup>. The genetic cost of an innate allele at step  $i$  is  $g_i$ .

If the set of innate steps is  $m$ , the probability of completing the entire chain is

$$P(m) = \prod_{i \notin m} q_i,$$

then the expected fitness is

$$W(m) = 1 + B \cdot P(m) - \sum_{i \notin m} c_i - \sum_{i \in m} g_i.$$

Consider now a rare mutant that makes an additional step  $k \notin m$  innate (changing the set to  $m' = m \cup \{k\}$ ), the change in expected fitness is

$$\Delta W_k = B \cdot P(m) \cdot \left( \frac{1}{q_k} - 1 \right) + c_k - g_k.$$

The first term is the marginal increase in completion probability, multiplied by the fitness benefit  $B$ . This depends on the product of non-innate learning probabilities for all remaining non-innate steps except step  $k$ . This term is larger when  $q_k$  is small (the step  $k$  is harder to learn). The second term is simply the saved learning cost. The third is the genetic cost that the mutant must overcome to spread.

Let us now compare two steps that are candidates for genetic assimilation,  $j, k \notin m$ . Step  $j$  is favored to be assimilated before step  $k$  if

$$\Delta W_j > \Delta W_k$$

which rearranges to

$$\left( \frac{1}{q_j} - 1 \right) - \left( \frac{1}{q_k} - 1 \right) > \frac{(c_k - g_k) - (c_j - g_j)}{B \cdot P(m)}$$

This condition explicitly shows that assimilation order depends on learning difficulty (left side) and differences in net cost (right side), scaled by the current state given by  $P(m)$  and  $B$ . That is, steps that are harder to learn and have higher learning costs offer higher marginal gains when made innate and so become assimilated earlier. High genetic costs oppose assimilation, delaying or preventing fixation of the mutant. But as more steps become innate, these change the landscape for the remaining steps, making the relative importance of cost differences become less important. The process of evolving instincts may then snowball. Because each fixation increases  $P(m)$ , the marginal term for the remaining steps rises, potentially converting previously neutral steps into favourable ones.

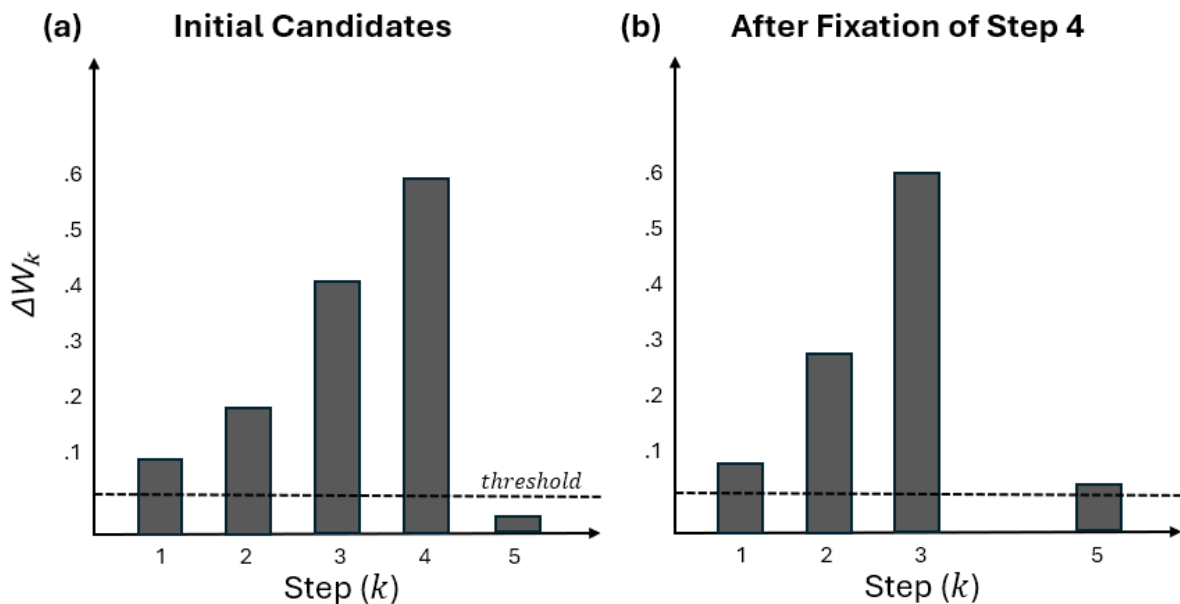
---

<sup>4</sup> For simplicity we treat learning difficulty ( $q_k$ ) and genetic encoding costs ( $g_k$ ) as independent across steps. Correlation between these quantities would primarily rescale marginal gains  $\Delta W_k$  and thus alter the ordering or speed of assimilation rather than the qualitative dynamics. Positive correlation would dampen assimilation by offsetting the benefits of hard-to-learn steps with higher encoding costs, whereas negative correlation would accelerate the snowballing pattern. The greedy, path-dependent structure of the process remains unchanged.

If assimilation proceeds iteratively along these lines, it may be well approximated by a greedy algorithm<sup>5</sup>:

1. Start with no innate steps ( $m = \emptyset$ )
2. Compute  $\Delta W_k$  for each step  $k \notin m$
3. If all  $\Delta W_k \leq 0$ , stop (no further assimilation is favoured)
4. Otherwise, fix the next step with the greatest positive  $\Delta W_k$  (assimilate it), updating  $m$
5. Repeat step 2 until no steps remain beneficial or all are innate

This process, as depicted in Figure 3, likely accelerates assimilation because each fixation raises  $P(m)$ , increasing the marginal term for remaining steps. The greedy approach, which picks the step with the largest immediate advantage, is biologically plausible and computationally efficient compared to evaluating all subsets, though global optimality is not guaranteed in cases with strong non-linearities or complementarity effects among steps. We note that the greediness is appropriate here given that Natural Selection is inherently “myopic” in that it climbs the steepest local gradient. By treating assimilation as a “greedy” algorithm, we model evolution as a process that prioritizes fixing the most expensive chokepoints in the hypercycle first.



**Figure 3. Greedy assimilation of heterogeneous behavioural steps.** Each bar shows the marginal fitness gain  $\Delta W_k$  from genetically assimilating a currently learned step  $k$ , computed from  $\Delta W_k = BP(m) \left( \frac{1}{q_k} - 1 \right) + c_k - g_k$ . Bars above the horizontal dashed line ( $\Delta W_k = 0$ ) represent steps for which innate encoding is selectively favoured. At any moment, selection fixes the step with the largest positive marginal gain (highlighted), approximating a greedy algorithm. Panel A: initial set of candidates with heterogeneous learning difficulties and costs; the hardest and most costly step yields the greatest benefit and is assimilated first. Panel B: after fixation of that step, the baseline completion probability  $P(m)$  rises, increasing the marginal benefit of assimilating each remaining step and shifting all bars upward. Previously neutral or weakly deleterious steps may therefore become favourable, illustrating how assimilation can accelerate and produce path-dependent “snowball” dynamics. This visualization emphasizes that heterogeneous steps compete by marginal value rather than evolving simultaneously.

<sup>5</sup> Greedy fixation need not be globally optimal when there are strong superadditive complementarities among particular pairs of steps; however, natural selection’s local, gradient-climbing nature makes the greedy approximation biologically apt in many contexts.

If learning ability  $q$  itself evolves, the invasion conditions become dynamic. Higher population-level learning ability raises the baseline  $P(m)$  and diminishes marginal gains from fixing more steps and so slowing assimilation. Conversely, early fixation of challenging steps can reduce selection to maintain high learning ability, creating positive feedback that accelerates further assimilation. The interplay between heterogeneous learning probabilities, per-step learning costs, and genetic costs can produce complex, path-dependent evolutionary trajectories<sup>6</sup>.

Heterogeneity among steps introduces a rich structure to assimilation dynamics. Steps that are hardest to learn and most costly to acquire tend to be assimilated first, but genetic costs and evolving learning ability can reverse or delay this pattern. Because each assimilation changes the payoff landscape, the sequence of fixation is path-dependent, perhaps producing snowball or cascading effects. This framework explains why behavioural hypercycles rarely evolve in a uniform or predictable order. It may also help explain why early events can lock populations into distinct evolutionary paths.

## 2.2. Learning and Instincts as Substitutes and Complements

In previous sections, we showed how a Good Trick implemented as a behavioural hypercycle can evolve either to remain learned or to become innate. We also saw that the capacity for learning (summarized by a per-step success probability  $q$ ) coevolves with the degree of genetic assimilation. Here we examine a deeper question: Does evolving a larger brain that learns more effectively necessarily reduce the evolution of instinct? Or can improved learning capacity *facilitate* more instinctive modules?

We define the following variables:  $K$  as total behavioural “brain capital,” reflecting the overall size or capacity of the nervous system,  $q(K)$  per-module learning success probability, increasing with  $K$  with diminishing returns,  $\beta(K)$ : innate economies-of-scale bonus, capturing integrative efficiencies from encoding multiple innate modules; non-decreasing in  $K$ ,  $\delta(K)$  as the genetic or developmental cost of innately encoding a module, typically decreasing with  $K$ ,  $\gamma$  as the residual cost or inefficiency associated with learned modules, and, finally,  $M(K)$  as total number of behavioural modules available, increasing linearly with  $K$  for simplicity. We model each behavioural module as either innate or learned (for simplicity, we normalize the base learned payoff to 1.). The net payoffs per module are:

$$\pi_I(K) = \beta_0 + \beta(K) - \delta(K) \text{ (innate)}$$

$$\pi_L(K) = q(K)(1 - \gamma) \text{ (learned)}$$

A module is genetically encoded when  $\pi_I(K) > \pi_L(K)$ . Thus, the total number of innate modules is

$$M_I(K) = M(K) \cdot 1_{\{\pi_I(K) > \pi_L(K)\}}$$

This structure highlights two opposing forces. One is a “substitution effect”: As learning becomes more reliable with increasing  $K$ ,  $q(K)$  rises, increasing  $\pi_L$  and reducing the evolutionary advantage of innate encoding. The other is an “income (capacity) effect”: Larger brains host more modules, gain greater innate integration bonuses  $\beta(K)$ , and face reduced innate costs  $\delta(K)$ , making innate encoding more advantageous. We can define the payoff gap as

$$\Delta(K) = \pi_I(K) - \pi_L(K).$$

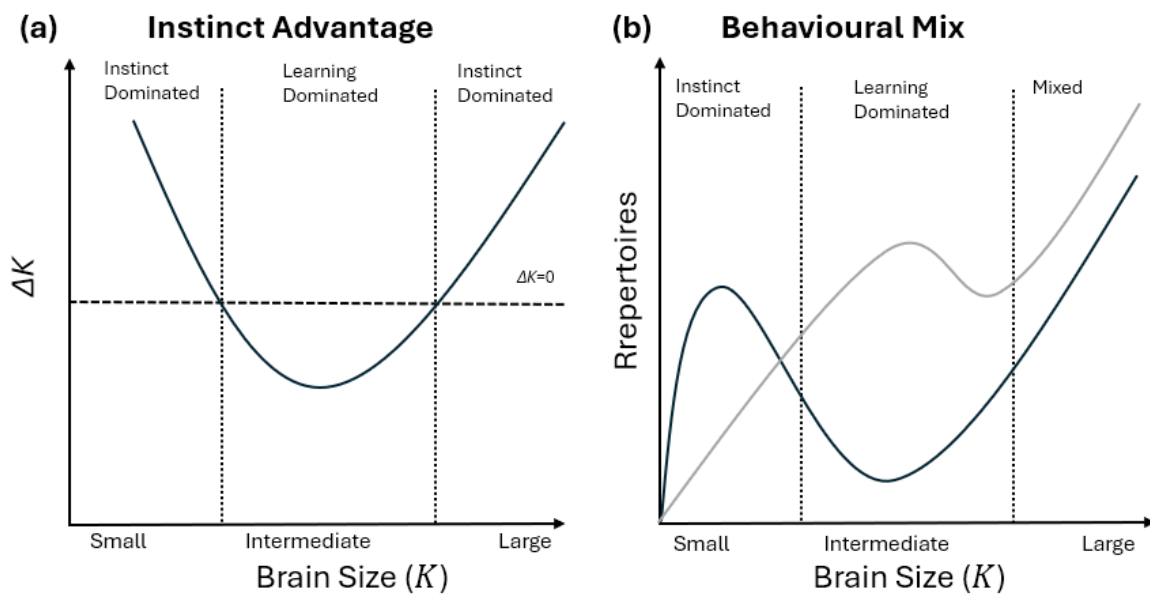
---

<sup>6</sup> Note that this formulation treats learning probabilities  $q_i$  as fixed. A realistic extension would incorporate that genetic assimilation of some steps frees cognitive capacity (attention, memory, practice time), potentially increasing  $q_j$  for remaining learned steps — a form of learning-instinct complementarity via automaticity. This positive feedback would accelerate subsequent assimilations and potentially alter optimal ordering, making hard-to-learn steps even more urgent early targets.

Instincts reappear at high brain capital if and only if  $\Delta(K)$  crosses zero twice, producing a U-shaped relationship between brain capital and reliance on innate behavioural modules. This happens when the derivative of the payoff gap changes in sign:

$$\frac{d}{dK} \Delta(K) = \frac{d}{dK} [\pi_I(K) - \pi_L(K)] = \beta'(K) - \delta'(K) - q'(K)(1 - \gamma).$$

As depicted in Figure 4a, if  $\Delta(K)$  is non-monotonic (positive at small  $K$ , negative at intermediate  $K$ , and positive again at large  $K$ ) then instincts re-emerge in large-brained organisms. Initially, learning improves rapidly ( $q'(K)$  is large), making learned strategies dominant. Later, as learning gains plateau, innate advantages grow faster from economies of scale and reduced genetic costs. That is, this pattern requires that  $q'(K)(1 - \gamma)$  initially exceeds  $\beta'(K) - \delta'(K)$ , but that this inequality reverses at higher  $K$  (i.e., learning efficacy initially improves faster than innate economies of scale, but that innate advantages eventually dominate as brain size increases).



**Figure 4. Brain size, instinct advantage, and behavioural repertoire.**

(a) Shows the payoff gap  $\Delta(K)$ , representing the “instinct advantage” of genetically encoding a behavioural module relative to learning it. Horizontal dashed line marks  $\Delta(K)=0$ , where innate and learned strategies are equally favoured. Regions of small, intermediate, and large brain size are indicated by vertical dotted lines. Positive  $\Delta(K)$  indicates modules for which innate encoding is selectively favoured, while negative  $\Delta(K)$  indicates modules dominated by learning. The U-shaped curve illustrates that instincts may dominate in small brains, give way to learning in intermediate brains, and re-emerge in large brains due to economies of scale and reduced genetic costs and as the payoff gap rises. (b) Realized behavioural repertoire across brain sizes. Black curve shows the number of innate modules, grey curve the number of learned modules. Vertical dotted lines mark the same small, intermediate, and large brain regions as in Panel A. In small brains, innate modules dominate; in intermediate brains, learning dominates; in large brains, both innate and learned modules coexist, reflecting the complementary effects of increased brain capacity and selective advantage of certain innate modules.

In other words, the key analytical condition is that the relative rates of change of the innate payoff components and learning efficacy determine the three regimes. Intermediate brains (where learning dominates) occur when the rate of increase of learning effectiveness  $q(K)$  initially exceeds the rates of increase in innate economies of scale  $\beta(K)$  and decrease in innate costs  $\delta(K)$ . Large brains (where instincts re-emerge) arise when, at higher brain capital, the rates of increase in innate benefits  $\beta(K)$  and reductions in genetic cost  $\delta(K)$  surpass the rate of increase in learning efficacy. This could happen when learning efficacy  $q(K)$  exhibits diminishing returns with increasing  $K$ , while innate advantages  $\beta(K)$  and  $\delta(K)$  continue to scale or improve. Thus, the balance of these

opposing rate-of-change trajectories governs evolutionary outcomes, perhaps producing the form of behavioural complexity observed across species.

In sum, learning and instinct are not simple substitutes and can function as complements over evolutionary time. As brain size grows, learning may initially displace instincts, but beyond a threshold, innate encoding becomes advantageous again due to economies of scale and reduced costs. This motivates the empirical prediction that large-brained organisms should exhibit both expanded learning capacity and expanded instinctive repertoires. How learning and instincts may be combined by such organisms is considered in the next section, which explores how changes in brain capital alter the distribution of assimilation across levels.

### 2.3. Compositional Structure and Differential Genetic Assimilation

The previous sections treated behaviour as a linear sequence of steps that could be acquired either through learning or genetic assimilation. In nature, however, complex behaviour is rarely a single indivisible sequence. Instead, adaptive solutions are built from reusable behavioural elements (motor acts, perceptual routines, or response patterns/schemas) that are combined into structured configurations such as chains, hierarchies, or parallel modules. Different tasks often draw on overlapping subsets of these elements, recombining them in task-specific ways. This compositional structure allows selection and learning to operate at multiple representational levels, rather than acting uniformly on behaviour as an indivisible whole. It also mirrors treatments of language evolution (e.g., Pinker & Jackendoff, 2005) where phonological primitives and syntactic structure are subject to distinct learning and genetic constraints.

A useful analogy comes from associative learning. An instrumental behavioural chain can be decomposed into (i) the elements themselves — the specific responses or skills — and (ii) the links that connect them, specifying when and how one element follows another. Learning can improve both levels: animals may refine the execution of individual responses, and they may also learn how to sequence, coordinate, or conditionally deploy those responses in novel contexts. Genetic assimilation, likewise, need not act on the entire behavioural chain at once. Instead, it can selectively encode either particular elements, particular linkages, or both.

Let behaviour comprise  $L$  representational levels indexed by  $\ell = 1, 2, \dots, L$ . Level  $\ell = 1$  corresponds to “primitives”, level  $\ell = 2$  to links, and higher levels to nested plans or hierarchical structure. Each level contains  $n_\ell$  components, of which  $k_\ell$  are genetically encoded and  $n_\ell - k_\ell$  are learned. Learning at each level is characterized by a probability of successful acquisition  $q_\ell \in [0, 1]$ , and a per-component learning cost  $c_\ell > 0$ . Components are assumed to be independent within and across levels (a conservative assumption; relaxing it strengthens our qualitative predictions)<sup>7</sup>.

Successful task completion requires that all components across all levels function correctly (whether learned or genetically encoded). Assuming acquisition across levels is independent, if all components are learned, the probability of successful performance is

$$Q_{\text{learn}} = \prod_{\ell=1}^L q_\ell^{n_\ell}$$

If  $k_\ell$  components at level  $\ell$  are genetically encoded, the probability becomes:

$$Q(k) = \prod_{\ell=1}^L q_\ell^{n_\ell - k_\ell}$$

---

<sup>7</sup> Note that this compositional model collapses to the sequential model in Section 2.1, if  $L = 1$  (no compositional links or higher levels), or if  $q_\ell = q$  and  $c_\ell = c$  for all  $\ell$ , or if no reuse advantage exists ( $\alpha = 0$ ). In those cases, the distinction between primitives and links is unnecessary, and assimilation proceeds on the entire behavioural chain as a single unit.

where  $\mathbf{k} = (k_1, \dots, k_L)$  and  $k_\ell \leq n_\ell$ . If  $k_\ell = n_\ell$ , level  $l$  contributes a multiplicatively neutral factor of 1. Because increasing  $k_\ell$  replaces one stochastic component (success probability  $q_\ell$ ) with a deterministic one, the marginal increase in overall success probability is  $\frac{\partial Q}{\partial k_\ell} = Q(\mathbf{k})(1 - q_\ell)$ .

Total learning cost is:

$$C(\mathbf{k}) = \sum_{\ell=1}^L (n_\ell - k_\ell)c_\ell$$

Costs are additive per component, whereas benefits combine multiplicatively because failure of any required component reduces overall task success. This reflects the hypercycle-like complementarity of behavioural elements, equivalent to the "O-ring production function" set up modelled by economists (Kremer, 1993). Primitives can be reused across tasks, whereas compositional links are typically task-specific (so reuse benefits may heavily favour assimilation of component behaviours but are likely small or effectively zero for links). To capture this, we define the effective benefit of genetically assimilating components at each level. Let  $B$  denote the baseline fitness benefit of successful task completion. We assume that the benefit of primitive assimilation is amplified by reuse:

$$B_\ell = \begin{cases} B(1 + \alpha R) & \text{if } \ell = 1 \\ B & \text{if } \ell > 1 \end{cases}$$

where  $R$  is the number of additional tasks that use the same primitive, and  $\alpha \geq 0$  is a constant that scales the advantage of reuse. Encoding components carries an opportunity cost, represented by  $g_\ell$  per genetically encoded component at level  $\ell$ . Links, being task-specific, rarely yield reuse benefits; primitives may be invoked across multiple hypercycles. The expected fitness is therefore

$$W(\mathbf{k}) = B \cdot Q(\mathbf{k}) - C(\mathbf{k}) - \sum_{\ell=1}^L k_\ell g_\ell.$$

For simplicity, the total task payoff is  $B$ ; reuse benefits enter only through marginal effects on selection for encoding particular components. That is, reuse does not increase the payoff of this task *per se* but increases the marginal value of encoding reusable components because those components contribute to additional tasks. Selection favours genetic assimilation at level  $\ell$  when the marginal fitness gain exceeds the genetic cost:

$$\frac{\partial W}{\partial k_\ell} > 0 \Leftrightarrow g_\ell < B_\ell Q(\mathbf{k})(1 - q_\ell) + c_\ell$$

This condition makes clear that selection for assimilation at a given representational level depends on three factors: learning difficulty  $(1 - q_\ell)$ , marginal learning cost  $c_\ell$ , and reuse-modulated benefit  $B_\ell$ . The term  $Q(\mathbf{k})(1 - q_\ell)$  is the expected increase in overall success probability from making component  $l$  deterministic rather than learned. A component at level  $l$  will be genetically assimilated when its expected contribution to fitness, accounting for difficulty, reuse, and learning costs, exceeds its genetic/developmental cost.

These expressions show that genetic assimilation acts preferentially on components with the greatest effective learning burden, defined jointly by learning difficulty  $(1 - q_\ell)$ , learning cost  $c_\ell$ , and reuse-modulated benefit  $B_\ell$ . This burden is conditional on the compositional structure of the task: a small number of difficult-to-learn links may be assimilated even when numerous easier primitives remain learned, or vice versa. This predicts the stable evolution of mixed systems, in which some behavioural components are genetically canalized while others may remain plastic. It explains why "low-level" routines (like the motor act of grasping) are almost always going to be innate across species: their "reuse value" is vast compared to the "link value" of a specific, one-off task. Innateness, on this view, is not a property of behaviours as wholes, but of representational levels within a behavioural architecture. Because primitives may be reused across multiple hypercycles while links are typically task-specific, the effective benefit associated with primitive assimilation can be substantially larger, even when per-unit learning costs are comparable. As a result, selection can favour innate primitives combined flexibly through learning, or alternatively innate compositional

scaffolds with learned fill-ins, depending on the relative learning difficulty and reuse value<sup>8</sup>. This suggests that in high-capacity systems, the hypercycle doesn't just grow longer; it becomes hierarchical, with instinctual modules serving as the reliable 'hardware' upon which plastic "software" can operate.

By making explicit the compositional structure of behaviour, this extension clarifies how learning and instinct can act as complements and provides a formal basis for understanding the evolution of behavioural modularity, hierarchical control, and innate scaffolding across a wide range of species and task domains. Changes in brain capital can shift not only the total balance between learning and instinct, but also the representational level at which genetic assimilation is favoured. Selection, moreover, could favour partial canalization: reusable primitives encoded innately; variable linkages left to learning. This demonstrates not only that components within a behavioural hypercycle can be expected to coevolve in mutually reinforcing ways but that learning and instinct evolve as complementary specializations across representational levels rather than as opposing alternatives. Consequently, the differential selective pressure across representational levels suggests a functional rationale for the anatomical segregation of learning and instinct within distinct neural circuits.

### 3. Discussion

The tidy intuition that the bigger the brain, the less hard-wired instinct, may be plausible at first glance, with cognitive diversity in nature scaling from tiny-brained robotic automatons on one end of a continuum to large-brained, thinking and learning machines on the other. But this is incomplete. The mechanism at the heart of our claim is the behavioural hypercycle: complex actions — building a nest, hunting with a sequence of moves, or making a tool — can be seen as teams of modular steps, where each can be learned or hard-wired. When the separate steps fit together as force multipliers of each other's effects, the whole chain is worth more than the sum of its parts where specialists cooperate, division of labour pays off, and the assembled sequence achieves disproportionate benefit. Given that economic structure, evolution faces an economizing choice: should a particular step be left to lifetime learning (e.g., paying the cost of practice each generation) or should it be genetically encoded (paying a developmental or pleiotropic cost instead)? We proposed that Natural Selection resolves that question one step at a time, "greedily" codifying the most expensive to learn steps first in the genome, with the answer for any step depending on both how many other steps are already innate and how good the species is at learning. This iterative trade-off (learning repairs errors and reduces the need for perfect genes while genes economize on costly learning) sculpts patterns of instinct and behavioural plasticity.

Supposing that learning and instinct co-evolve in this way we found that the relationship between neural capacity or "brain capital" and the balance of instinct versus learning is not a simple inverse but curvilinear. Three broad evolutionary regimes are predicted as brain capital increases. First, when brains are small, instincts dominate: animals are built like precise machines, with tight, canalized routines that are reliable and cheap to run. Second, with more neural capacity, selection often favours learning: organisms become flexible, offloading many behaviours to ontogenetic acquisition rather than fixed genes. This is the "substitution" regime — learning replaces brittle instincts because it is cheaper and more adaptable. Third, beyond a further threshold of capacity the trend can reverse: instinct re-emerges. Large-scale nervous systems (or large, highly integrated collectives) can support many specialised, hard-wired modules while still maintaining high learning ability.

Why does this "instinct re-emergence" happen? Again, think in economic terms. Small brains rely on instincts because they cannot afford the brain capital required to support flexible learning; intermediate brains invest in learning to better bank on brain power by relying less on instincts. But

---

<sup>8</sup> Such as an innate 'grammar' of nest building, where the overarching structure is hard-wired but the specific 'fill-ins' (materials used) are learned according to local availability.

once brain capital is vast enough, two effects appear. The “substitution effect” still exists (i.e., learning can replace particular instincts) but an “income effect” now operates meaning that additional capacity increases the feasible stock of innate modules. This increased capacity raises the ceiling of what the organism can afford to store and coordinate. Encoding a set of specialised modules may become relatively cheaper per unit (economies of scale), while integrating many innate pieces unlocks new synergies that learning alone cannot cheaply or reliably reproduce. In other words, big brains can “buy” both flexibility and an expanded catalogue of ready-made routines and the payoff from making some behaviours innate rises with system size. Depending on whether this substitution or income effect predominates, the evolution of greater brain size may increase the degree to which animals rely on learning versus instinct respectively. This has ramifications for when to expect learning-dominated species, when to expect instinct-rich ones, and when both will coexist.

Our final extension suggests that this trade-off does not operate only at the level of whole behaviours. This predicts that larger or more cognitively complex organisms should not simply possess more learning or more instinct, but increasingly hierarchical mixtures of both: innate building blocks combined through learned composition. Many real behavioural repertoires in cognitively complex animals appear highly compositional: complex actions assembled from reusable primitives, linked together and organized in higher-order plans. Selection and learning therefore operate at multiple representational levels, not merely on whole behaviours but on the reusable primitives from which they are composed. If so, learning and instinct emerge not merely as substitutes but as complements. Components that are reused across many tasks or that are especially expensive or unreliable to learn are predicted to become genetically canalized, whereas task-specific linkages or flexible combinations are more likely to remain plastic. In this sense, evolutionary increases in brain capital may shift not only how much behaviour is learned, but at what level learning operates.

In short, our analysis suggests that instinct and learning are complementary specializations within a shared behavioural architecture. Evolution does not choose between “genes” and “experience” wholesale but allocates control across components and representational levels, canalizing those elements that are costly or widely reused, while leaving flexible the links and combinations that benefit from context sensitivity. Innateness is therefore expected to concentrate in reusable primitives and structural scaffolds, whereas learning should dominate task-specific sequencing and recombination that repurposes those primitives. This perspective points to several testable predictions. First, behaviours that are reused across many ecological contexts or tasks should be disproportionately innate relative to equally important but context-specific behaviours. Second, across species, increases in brain capital should be associated not simply with more learning, but with more hierarchical mixtures: expanding repertoires of innate building blocks coordinated by flexible learning. Third, evolutionary or artificial selection experiments should preferentially assimilate the hardest-to-learn or most failure-prone components of a task rather than entire behavioural chains at once. Finally, developmental and neural analyses should reveal that the circuitry supporting learning also marks the loci at which genetic canalization accumulates over evolutionary time. In this way, mechanisms of learning provide a map of likely evolutionary construction, and the organization of instinctive behaviour records the learning problems that ancestral organisms solved often and reliably enough to encode into their genomes.

### *3.1. Evidence and Implications*

This section outlines empirical foundations supporting the biological plausibility of the model. Namely, we assume that (i) animals spontaneously generate novel combinations of behavioural components, (ii) neural circuitry can specify phenotypes with sufficient precision to canalize behavioural components or their linkages, and (iii) behavioural control systems are hierarchically organized in larger brains in ways that can exploit hypercyclic organization. Together, these properties provide the raw material and mechanisms required for the formation and genetic assimilation of behavioural hypercycles.

### 3.1.1. Hypercycle Generation: Recombination and Behavioural Spillover

For a hypercycle to be learned, organisms must first generate multi-step combinations of existing behavioural elements. Observations from across ethological and learning literatures suggests that opportunities for such recombination arise naturally and frequently. Classic “displacement activities” are particularly demonstrative that heightened levels of motivation can trigger the expression of behavioural modules outside their usual context (Tinbergen, 1951). For instance, territorial sticklebacks may suddenly interrupt fighting to engage in nest-digging behaviours (Van Iersel & Tinbergen, 1948). Some types of birds similarly interrupt combat to suddenly engage in various types of nesting or food-pecking behaviours, while sexual frustration also appears to lead to all sorts of displaced behaviours including food-begging, preening, food-catching movements, food-pecking movements (Tinbergen, 1951). These cases indicate that innate components can be activated alongside other behaviours in novel or unusual situations when animals face uncertainty or competing drives, potentially making them subject to modification and repurposing by reinforcement-based learning should such combinations prove serendipitous. Rodents, for example, co-opt substrate-moving routines normally used for food caching or nest building into defensive burying behaviour directed at threatening or unfamiliar stimuli (De Boer & Koolhaas, 2003). This behaviour is flexibly calibrated by experience and can be integrated with other defensive acts, illustrating how previously independent components become organized into functional chains. Similar redeployments occur in the wild, such as in the case of California ground squirrels that combine substrate throwing with evasive manoeuvres to disrupt rattlesnake predation (Owings & Coss, 1977; Rundis *et al.*, 2007). These examples show that existing routines can be recombined into coordinated multi-step solutions whose benefits depend on complementarity among parts. Further evidence that behavioural primitives are modular and can be recombined is provided by observations of domestic dogs. Predatory sequences in wolves (e.g., orienting → stalking → chasing → biting) appear as separable components have been selectively reorganized through training and breeding into specialized tasks such as herding (McConnell, 1990). Genetic studies implicating axon-guidance and neurodevelopmental genes in breed-specific behavioural repertoires suggest that such reorganized chains can subsequently become canalized (Dutrow *et al.*, 2022; Jeong *et al.*, 2025), consistent with our discussion below that such genes provide a plausible basis for the genetic assimilation of hypercycles.

Numerous laboratory observations further support the same general point. Even in species with relatively simple nervous systems, such as fruit flies, individuals will often spontaneously generate behaviour under certain conditions (Maye *et al.*, 2007). Under frustrative non-reward or changing contingencies, rats show “resurgence” of previously learned behaviours and often combine them in novel ways (Winterbauer & Bouton, 2010). In a classic example, pigeons trained separately to push a box and to climb it spontaneously combined these behaviours to reach an otherwise inaccessible reward (Epstein, 1987), reminiscent of similar combinatory problem solving as observed in crows, where previous behaviours are combined to achieve reward in response to novel challenges (Weir *et al.*, 2002). Novelty potentiates arousal and mesolimbic dopamine transmission, which tend to generate spontaneous movements at moderate levels and, at high levels (such as likely corresponding to intense reinforcement in nature) lead to the tight integration of such movements into fixed or stereotyped and automatic patterns (Lubow, 1989).

Across a range of taxa so far studied, behavioural variability, exploration, and associative learning processes therefore appear conducive to the generation of ordered sequences from pre-existing parts. A proximal mechanism for this process in species located along the higher end of brain capital spectrum may lie in the systems underpinning Pavlovian-Instrumental Transfer (PIT). This refers to the acute intensification of instrumental responding when animals are presented with Pavlovian stimuli that have previously been conditioned in association with related outcomes (e.g., food rewards) (e.g., see Corbit & Balleine, 2005). In nature, PIT could plausibly function as a powerful search engine for candidate hypercycles. By allowing previously conditioned stimuli to invigorate instrumental actions, the Basolateral Amygdala (BLA) and Nucleus Accumbens (NAc) effectively

gate exploration towards “plausible” links, i.e., those involving stimuli already associated with reinforcement. This Pavlovian “invigoration” increases the probability that an animal will discover the synergistic benefits of a new behavioural chain. Furthermore, mechanisms of instrumental incentive learning involving the insular cortex ensure that these emerging chains are stabilized only when they are relevant to the organism's internal motivational state (Balleine & Dickinson, 1998). In short, it is tempting to suppose that these circuits provide the behavioural scaffold upon which selection operates: by temporarily stabilizing a learned link through PIT and incentive gating, these systems provide the consistent performance required for genetic assimilation to “lock in” the connection over evolutionary time.

These observations collectively suggest the antecedent conditions for our model exist in nature. Animals may naturally produce and stabilize structured combinations of complementary actions under potentially high stakes conditions (novelty, frustration, highly motivated and competitive interactions) where Good Tricks are especially likely to be useful. Such spontaneously assembled chains satisfy the formal requirements of the “proto-hypercycles” assumed in our model, providing abundant substrates upon which selection can subsequently act through genetic assimilation. Importantly, the associative learning mechanisms that enable chaining are phylogenetically widespread, including in arthropods (e.g., Bitterman *et al.*, 1983; Hazlett, 2007) and molluscs (Sahley *et al.*, 1981) indicating that the capacity to generate such combinations is not restricted to a few unusual species but is a general feature of complex behaviour.

### 3.1.2. Genetic Specificity in Neural Wiring

Another assumption of our model is that selection can canalize particular behavioural links or subroutines (rather than only coarse traits such as overall brain size, temperament, or general learning ability). This is because genetic assimilation in our model is mostly interpreted not as the encoding of entire complex behaviours, but as either or both the progressive stabilization of specific circuit components or their intervening links within an otherwise plastic hierarchy. This requires that the genome possess sufficient “addressing power” to specify connectivity at the level of identifiable circuit elements. Neurodevelopmental genetics strongly supports this premise. Since Sperry (1963) first proposed that neurons bear molecular identity tags that guide partner selection during development, extensive evidence has accumulated for concrete molecular substrates for this idea (Jin & Lee, 2019). For example, in vertebrates, clustered protocadherins (Pcdhs) generate combinatorial cell-surface “barcodes” through stochastic promoter choice, producing unique identity signatures for individual neurons (Chen & Maniatis, 2013; Wu & Maniatis, 1999). In insects, the single *Dscam1* locus achieves comparable combinatorial diversity through mutually exclusive alternative splicing, generating tens of thousands of isoforms (Jin & Lee, 2019; Wu & Maniatis, 1999; Yue *et al.*, 2016). Despite their distinct genomic architectures, both systems implement the same functional principle: highly specific homophilic binding that enables neurons to distinguish self from non-self and wire selectively to appropriate partners (Wu & Maniatis, 1999). The recurring theme across vertebrates, arthropods, and independently in cephalopods, where large expansions of protocadherins have evolved convergently (Styfals *et al.*, 2019), suggests that fine-grained molecular “addressing” is a common solution to building large nervous systems.

Importantly for our purposes, these mechanisms are required for precise circuit assembly. Manipulations of protocadherin or *Dscam* diversity disrupt dendritic tiling, axonal targeting, and stereotyped connectivity patterns, leading to predictable defects in neural architecture and behaviour (Dong *et al.*, 2025; Wu & Maniatis, 1999). Across taxa, many species-typical actions (escape reflexes, orienting responses, courtship displays, and coordinated motor programs) depend on genetically specified circuit motifs assembled prior to experience (Barabási *et al.*, 2025; Shen *et al.*, 2014). In mammals, molecularly defined spinal *motor synergy encoder* neurons and genetically specified hypothalamic ensembles orchestrate coordinated movements and instinctive behaviours (Azim *et al.*, 2014; Kiehn, 2016; Stagkourakis *et al.*, 2023). These findings demonstrate that behaviourally meaningful modules can, in principle, be developmentally hard-wired.

At the same time, the molecular boundary between innate and learned behaviours appears surprisingly porous. Major transcriptional pathways that organize circuits during development (e.g., CREB-dependent gene expression and neurotrophin signalling) are also later reused for long-term potentiation and memory consolidation in the adult brain (Chowdhury, An, & Jeong, 2023; McDowell *et al.*, 2011; Sakamoto, Karelina, & Obrietan, 2011), with disruptions to these shared pathways impairing both developmental circuit stabilization and long-term memory (McDowell *et al.*, 2011). This means that the same biochemical machinery that establishes initial (and therefore, potentially innate) connectivity also supports experience-dependent refinement. Such continuity provides a plausible mechanistic substrate for Baldwin-type processes: circuits first assembled through plasticity can, under consistent selection, become developmentally stabilized or “fixed” via allelic substitution among genetic variants affecting the same molecular toolkit (Barabási *et al.*, 2025; Bavelier *et al.*, 2010).

In sum, the convergent evolution of high-diversity recognition molecules across insects, vertebrates, chelicerates, and cephalopods (Jin & Lee, 2019; Styfhals *et al.*, 2019), the demonstrated necessity of these systems for precise circuit formation (Wu & Maniatis, 1999), and the reuse of developmental plasticity pathways in adult learning (McDowell *et al.*, 2011) indicate that the genome is capable of specifying neural structure at fine levels of resolution. Selection can therefore plausibly act on individual links, primitives, or subroutines within behavioural assemblies, as assumed by our hypercycle framework.

### 3.1.3. Hierarchical Composition of Control

An assumption of Section 2.4 is that behaviour is organized hierarchically with relatively fixed motor primitives that are composed and sequenced by more flexible supervisory systems for the purpose of achieving challenging tasks. Convergent evidence from cephalopods, vertebrates, and even robotics corroborates this structure. A clear biological illustration is the octopus. Roughly two-thirds of its neurons reside in the arms rather than the central brain, and each arm contains semi-autonomous circuitry capable of executing complex grasping and exploratory routines locally. The central brain appears to specify high-level goals or intentions (e.g., locomote, manipulate, investigate), while peripheral circuits implement the detailed motor solutions using reflexive and context-sensitive subroutines (Sumbre *et al.*, 2001). This division of labour reduces the computational burden associated with controlling a complex soft body and closely mirrors “subsumption architectures” in robotics, where low-level reactive modules handle details of execution while higher layers select and sequence behaviours (Prescott *et al.*, 1999). In effect, the brain chooses *what* to do and the limbs determine *how* to do it<sup>9</sup>.

A comparable layering is evident in vertebrates. The dorsal striatum functions as a repository for “chunked” motor routines and habits, while medial temporal structures (e.g., the septo-hippocampal system) support flexible, relational control. Across rodents and primates, the dorsolateral striatum is preferentially involved in stimulus-response learning and automatic

---

The “Octopus Principle” of distributed control (innate primitives governed by learned links) described in our model enjoys a sort of “existence proof” in modern robotics. Early AI attempted to solve movement through “General Intelligence” (brute-force calculation of every joint angle), which proved too slow and brittle, while contemporary robotics has moved toward Subsumption Architecture and Morphological Computation, where “intelligence” is offloaded to the physical design of the limb (innate primitives) while the central processor focuses on high-level sequencing (learned links) (Brooks, 1991). The morphological computation principle is an especially good example of an “innate primitive”, in that the “intelligence” is built into the physical mechanics of the robotic limb so that it “knows” how to move automatically (Pfeifer & Bongard, 2006). This is very much the same sort of decomposition into primitives and linking control that our model formalizes. That engineers are converging on the same hierarchical solution as octopuses and primates suggests that our model describes a universal logic applicable to both carbon and silicon-based systems.

sequence execution, whereas more medial and anterior regions support goal-directed and wider associative functions (Gahnstrom & Spiers, 2020; Graybiel, 2008; Mattfeld & Stark, 2015; Yin & Knowlton, 2006). With practice, behavioural combinations can transition from deliberative control to striatal “chunking,” in which discrete movements are fused into unified, automatically triggered units. This process produces stable, reusable motor primitives that can be rapidly deployed with minimal cognitive oversight, and so may be an ontogenetic precursor to genetic assimilation. Studies suggest that these primitives are implemented as “motor synergies” or modules, that is, coordinated patterns of muscle activation that serve as the building blocks of movement (Flash & Hochner, 2005; Giszter, 2015). Such synergies are observed across frogs, rodents, and primates and appear highly canalized developmentally. In rats, for instance, many spinally controlled primitives are established by early postnatal life and remain remarkably stable even when normal motor experience is prevented, indicating that their structure is largely genetically specified (Yang, Logan, & Giszter, 2019). These findings imply that a substantial fraction of their motor repertoire is preassembled and robust to environmental variation (precisely the sort of substrate expected if selection can genetically assimilate frequently used behavioural components).

In contrast, the hippocampal system specializes in flexible linking and contextual control. It is required for configural learning, spatial and relational associations, and in the arbitration of competing behavioural tendencies (Gray & McNaughton, 2000; Mattfeld & Stark, 2015). Gray and McNaughton (2000) argue that the hippocampus functions as a “comparator” that gates which associations are expressed within a given context, using theta-modulated inhibition to protect selected representations from interference and to interrupt prepotent habits when novelty or conflict is detected. This contextual gating can be seen as compatible with the goal-directed control described by Balleine and colleagues. While the dorsolateral striatum stores “habitual” chunks, the dorsomedial prefrontal cortex (dmPFC) and its associated loops sequence and reorganize these components into novel chains based on current values (Ostlund, Winterbauer, & Balleine, 2009). The hippocampal comparator effectively “permits” this goal-directed system to take the wheel when the environment changes. Crucially, these complex sequences can themselves be “compiled” into unified automatic units through a process of action chunking, improving performance while remaining under high-level supervisory control. This means that the chains may be sequentially rigid but remain goal-directed at their point of initiation (Dezfouli & Balleine, 2012, 2013). This creates a nested hierarchy where the brain does not just learn a chain; it learns to treat an entire chain as a single, reusable “macro-primitive.” Thus, while the striatum stores “chunks,” the hippocampus manages their sequencing and recombination.

All these observations converge on a common principle corresponding to our predicted Regime III. Behaviour is constructed compositionally from stable, partly innate motor modules that are flexibly sequenced by higher-order systems. Lower levels provide reliable execution; higher levels provide combinatorial control. This organization is the substrate required by our model. Genetic assimilation should preferentially act on frequently reused primitives — stabilizing their implementation — while leaving the relational and sequencing layers comparatively plastic. The result is a hierarchy in which evolution progressively “hardens” the building blocks of behaviour while preserving flexibility in how those blocks are combined. The recurrence of this design in cephalopods, vertebrates, and engineered robotic systems suggests that it is not taxon-specific but a general solution to the adaptive control of complex action.

### 3.2. Predictions

Our model generates empirical predictions that diverge from canonical formulations of the Baldwin effect and from “general intelligence replaces instinct” views of brain evolution, making the framework falsifiable. If largely substantiated, however, the heights of cognitive evolution may not properly be seen as liberation from instinct, but the mastery of it. That is, the mark of adaptive cognitive complexity may not be the sort of high-powered correlational capacities of, say, Large

Language Model AI, but the ability to coordinate and appropriately deploy a massive library of species-specific subroutines within a plastic architecture.

### 3.2.1. Learning as Scaffold, not a Substitute

Classical views assume increasing learning capacity reduces the need for innate structure (e.g., MacPhail, 1982; Stephens, 1991) or is even predicated on the loss of it (Deacon, 1998). On the contrary, we predict that learning can *facilitate* the evolution of additional instincts by expanding the search space of discoverable behaviours. Moreover, in principle, any taxa capable of learning a class of behaviours should also be capable of evolving those same behaviours as instincts. Furthermore, lineages with greater learning capacities should, over evolutionary time, accumulate *more* innate modules rather than fewer<sup>10</sup>. Additionally, neural pathways that support rapid learning should overlap anatomically and molecularly with regions where instinctive routines are encoded, and genes implicated in instinct evolution should be expressed in circuits that previously supported learning.

A related prediction arises from our speculation regarding the role of the systems that underpin PIT in mammals. We suggested that this functions to facilitate the discovery of complementarity among behavioural responses, with the evolution of this capacity perhaps favoured because it enhanced the evolvability of instincts through the discovery and stabilization of hypercycles. If correct, PIT should have evolved also in analogous neural substrates of other groups with extreme levels of brain capital, such as octopuses. Speculating further, the observed distinction between General PIT (motivational invigoration) and Outcome-Specific PIT (sensory-specific direction) may functionally correspond to the discovery of “Primitives” and “Links” respectively. However, PIT tests are almost invariably conducted under extinction conditions, and so the longer-term consequence of PIT effects on learning remain unexplored. We propose designing an experiment where a Pavlovian cue is introduced not to see if it increases an *existing* instrumental response, but to see if it facilitates the *initial learning* of a complex, multi-step chain where the steps are individually weak but collectively strong. Similarly, in environments where behavioural hypercycles can be discovered through exploration and experience, animals with impaired BLA/NAc connectivity should be less capable of learning hypercycles than those without.

### 3.2.2. The Re-Emergence of Instinct at Large Brain Size

Whereas the idea that brain capital increases simply by evolution of ever greater intelligence predicts a *dilution* of instinct, our model predicts an *expansion* of specialized circuitry. This sees large brains more like “Swiss Army knives” of specialized modules, not blank-slate learners (Pinker, 2002). This puts the model firmly in support of the general working assumptions of Evolutionary Psychology. Along with standard narratives that predict a monotonic decline of instinct with increasing encephalization, and with larger brains imply fewer fixed behaviours, we expect that, as brain capital increases from small to intermediate values, learning substitutes for instinct. But beyond a threshold, we predict increased computational and representational capacity enables the discovery and subsequent canalization of increasingly complex hypercycles, and therefore an expansion of instincts. Further empirical work can therefore test the following predictions:

1. A U-shaped or mixed relationship between brain size and proportion of instinctive behaviour.
2. Large-brained taxa (e.g., corvids, parrots, cetaceans, primates, cephalopods) should exhibit both high learning ability *and* extensive suites of species-typical routines.
3. These routines should appear modular and recombinable rather than globally rigid.
4. Comparative neuroanatomy should show increased structural modularity and specialized circuits rather than uniform “general intelligence” scaling.

---

<sup>10</sup> Supportive evidence for this “income effect” can be found already for mammals (see Changizi, 2003).

### 3.2.3. Bottleneck-First Assimilation

Innateness should be unevenly distributed within behaviours, concentrated at reliability bottlenecks. Classical canalization accounts often treat behaviours as wholes. While this may be entirely appropriate for behaviours learned as if by the shaping procedure, our hypercycle model, which focuses more on the chaining procedure analogy, predicts *component-wise* assimilation. Selection should preferentially fix the steps that are most costly, slow, or failure-prone, leaving easier components plastic. This means we expect complex behavioural sequences should display partial, mosaic innateness, with the most deeply canalized elements corresponding to the hardest or most dangerous steps to learn. Paradoxically, in development, the easiest steps should remain experience-dependent longest. Closely related species without the same behavioural pattern (but that may be capable of learning something similar experimentally), should differ most strongly in their capacity at these bottleneck components.

A variety of examples that may be consistent with this logic includes (1) innate song templates with learned dialect variation; (2) predator or brood-parasite recognition biases with flexible tactics; (3) domestication targeting fear/stress bottlenecks; and (4) language showing innate structural constraints with culturally learned surface forms. Although there is recent evidence of anatomically distributed but developmentally scaffolded neural ensembles for instinct (Stagkourakis et al., 2023) indicating that species-typical routines are implemented as sophisticated circuit motifs (providing a plausible target for the component-wise genetic assimilation predicted by our model), more research is needed.

### 3.2.3. Ontogeny Partially Recapitulates Assimilation Order

Although we do not advocate a strong “ontogeny recapitulates phylogeny position”, we might expect skill automatization within individuals to mirror evolutionary genetic assimilation across generations. The learning processes of habit formation in individuals should employ the same neural transitions (e.g., chunking in the striatum) that genetic assimilation utilizes over time, and developmental trajectories should echo historical difficulty gradients. To the degree that this is correct, behavioural components closer to consummatory outcomes or with simpler contingencies should stabilize earlier in development (more distal, relational, or sequencing components should remain plastic longer). Development should often proceed from flexible assembly toward progressive automatization (“compilation”) of frequently used subroutines.

Cursory observations suggest instincts indeed mature in this way (Thorpe, 1956). For instance, inexperienced ravens appear to be very unselective in the materials they choose to construct nests – although the nest-building behaviours that follow material-gathering behaviours are displayed by these individuals closely resemble the behaviours of experienced adults. Over time, they learn to choose materials more discriminately to build more effective nests (Tinbergen, 1951). The hunting FAP of the spider wasp, *Pepsis cerberus* is also suggestive: Punzo (2005) describes how behaviours nearest to the beginning of the hunting sequence are initially more flexible. With repeated trials, efficiency improves through adjustments of the behavioural components nearest the beginning of the sequence (such as the initial approach and paralysation behaviours). By contrast, no change occurs in behaviours towards the end of the hunting sequence (such as burial and oviposition of the host). Another suggestive observation is that, among species of large felines, mothers often provide cubs with still-alive prey items – presumably to provide the cubs with opportunities to practice the killing components of hunting behaviour (Eaton, 1968). Aside from the obvious benefit to practicing such particular skills, the present thesis offers an additional interpretation: the mother is acting to facilitate learning in her offspring by giving behavioural components nearest to the kill behaviour primary reinforcement, thus making behaviours antecedent to the kill acquire greater salience for later learning through secondary reinforcement.

Evidence also comes from the studies of a hunting Fixed Action Pattern in ants. Among the species *Myrmica rubra*, there is variation among individuals in how innate their sequence of hunting behaviours is. A behavioural hypercycle explanation would predict that when the behaviour has not

become entirely innate, actions closer to the end of the sequence should be relatively more fixed, while those earlier in the sequence may still require reinforcement-learning. Among those ants that appear to require some learning, then, the behavioural components nearer the beginning of the chain should be more variable. This is in fact the case (Reznikova *et al.*, 2012). This is consistent with the idea that selection should stabilize behaviours nearer the “consummatory” end of a chain first, as these components provide the immediate reinforcement necessary to ground the learning of antecedent steps.

### 3.2.3. Combinatory Exaptation

Behavioural evolution should show exaptation and modular reuse rather than independent origin of each routine. If behaviours evolve through the recombination of existing modules, novel instincts should rarely arise from entirely new circuitry. We predict that distinct, complex behaviours should share reusable subcomponents, that there ought to be genetic and neural overlap between seemingly unrelated skills, and convergent behaviours across taxa should reuse homologous skill/primitives rather than independent solutions to adaptive challenges.

### 3.3. Limitations and Future Directions

While our model provides a clear basis for the re-emergence of instinct, its current form relies on several simplifying assumptions that invite further refinement. First, we assumed independence of components. In Section 2.1, we modelled the success probability of each step as independent. In nature, the acquisition of behavioural steps may often be synergistic; failure in an early step may preclude the execution of later ones, or learning one step might make another easier to acquire. Empirical studies of behavioural repertoires seem to show that behaviours covary in modular fashion (Werkhoven *et al.*, 2021), presenting a challenge to evolutionary models that assume independent traits (Wagner & Altenberg, 1996). However, we note that incorporating these non-linear dependencies would likely amplify the “snowball effect” we identified (§2.2), as the value of fixing one bottleneck becomes even higher when it “unlocks” the reliability of the entire chain. Moreover, if steps are interdependent, the “bottlenecks” become more critical, and reinforces our “bottleneck first” prediction.

Second, we treated “Brain Capital” as a static resource pool. A more dynamic model might account for the Metabolic Trade-offs inherent in encephalization. If the metabolic cost of building a larger brain exceeds the fitness gains of the new “Good Tricks” it can learn, the “Income Effect” we described may be suppressed. Future work could integrate the Expensive Tissue Hypothesis (Aiello & Wheeler, 1995), which highlights how larger brains are metabolically costly and must be offset by corresponding ecological gains or reductions elsewhere (e.g., gut tissue). This could help predict the specific ecological conditions (e.g., nutrient-dense diets) required to push a lineage into the Regime III “instinct re-emergence” zone.

Another limitation relates to the scope of behavioural domain and environment. We assume a relatively stable task structure and fixed external environment, whereas in nature environments may shift in ways that disrupt conditions needed for behavioural hypercycles to evolve. Finally, many key parameters (e.g., the cost of genetic encoding, learning ability, reuse effects) are hard to estimate and operationalize. The model is therefore best viewed as a qualitative, generative theory than a source of precise quantitative predictions.

#### 3.3.1. The “Automaticity” Force Multiplier

An interestingly speculative direction is related to the notion of “evolutionary automaticity”. Our model parallels neuroevolutionary arguments that selection for transferring learned control to subcortical habit systems has played a role in driving cognitive evolution (Shine & Shine, 2014), proposing a proximal mechanism that parallels genetic assimilation across longer timescales. Our model treats “innate” and “learned” as distinct categories, but neurobiology suggests a continuum.

Within a single lifetime, repeated learning transforms flexible, cortical "links" into rigid, sub-cortical "habits" (Balleine & O'Doherty, 2010) a process known as "automatization" or "chunking" (Graybiel, 2008; Poldrack *et al.*, 2005). As mentioned above, it is worth testing whether the same neural structures (e.g., the transition from the associative to the sensorimotor striatum) that handle lifetime automaticity are the primary sites for long-term genetic assimilation. If so, "automaticity" acts as a proximal simulation of evolution: by turning learned links into "pseudo-primitives," the brain frees up higher-level capital to coordinate even more complex chains. This creates a nested hypercycle where lifetime learning and evolutionary encoding could reinforce one another at different timescales.

### 3.3.2. Extended Phenotypes

While the present thesis focuses on sequential behaviours operating within a single organism, the same logic may be applied to phenotypes contributed jointly by genes acting across multiple bodies. In social species, for instance, the "links" and "primitives" of a behavioural hypercycle might be distributed across multiple related individuals, with the supervisory system acting at the "colony" or "household" level as individual members execute modular steps in a collective chain. To illustrate, we find a particularly compelling bridge to this idea in Becker's (1974) "Rotten Kid Theorem". We use this idea to argue in Appendix B that a "head of household" (the parent) closes the loop of social interactions by redistributing resources, effectively aligning the incentives of selfish members so they act in the interest of the collective hypercycle. This social coordination may initially be discovered through complex learning and cultural transmission. However, our model predicts that evolution should eventually assimilate the reliability bottlenecks of this social process as well. For example, if the success of the "Rotten Kid" logic hinges on the parent's ability to detect and correct selfish behaviour. If this is a task that is cognitively demanding and slow to learn, selection may favour the genetic canalization of specific social monitoring and altruistic "primitives." This creates a hybrid social architecture where the "hard parts" of cooperation become innate, while the specific tactical "links" of the social chain remain culturally plastic, raising potentially interesting connections to evolutionary theory (such as related to ideas of Haig (2001) that further work should explore).

## 4. Summary and Conclusions

We began with a tension between the clever generalist and the instinctive specialist. By formalizing behaviour as compositional hypercycles built by linking reusable behaviours, we saw that these need not be opposing evolutionary endpoints but emerge as outcomes of the same underlying process. This perspective sees complex behaviour as hierarchically organized and compositional. Behavioural sequences function like modular systems: semi-independent components can be added, duplicated, and specialized while higher-level control remains flexible. In this respect, the evolution of behaviour is reminiscent of the evolution of segmented anatomy, much like the Hox-gene logic of anatomical evolution, where complexity arises through the iterative refinement and recombination of reusable parts rather than wholesale innovation. Learning can initially substitute for brittle instinct, but as brain size ( $K$ ) accumulates and behavioural components are reused across contexts, economies of scale reverse this relationship, with selection favouring the genetic hard-wiring of hard-to-learn components. The result is not the replacement of instinct but its re-emergence in a richer, more modular form. Large brains thus contain libraries of calcified primitives dynamically orchestrated by clever control processes rather than operating as either a vast blank slate or rigid reflex machine.

From Katydid to Killer Whale, clear empirical consequences follow. Species with greater learning capacity should ultimately evolve *more*, not fewer, instinctive modules. Genetic assimilation should proceed mosaic-like, targeting the most failure-prone or costly tesserae first. Neural circuits that support rapid learning and automatization should also be the primary substrates of evolved instincts. Highly encephalized lineages should exhibit expanded repertoires of partially innate, recompilable behavioural motifs rather than a simple replacement of instinct by general intelligence.

If along the right lines, this economic logic combined with the hypercycle concept provides a account of how flexible cognition and specialized instinct can scale together, illuminating the shared design principles underlying complex agency across biological and even artificial systems.

## References

1. Aiello, L. C., & Wheeler, P. (1995). The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. *Current anthropology*, 36(2), 199-221.
2. Amsel, A. (1992). *Frustration Theory*. Cambridge: Cambridge University Press.
3. Amsel, A. (1994). Précis of Frustration theory: An analysis of dispositional learning and memory. *Psychonomic Bulletin & Review*, 1, 280-296.
4. Azim, E., Jiang, J., Alstermark, B., & Jessell, T. M. (2014). Skilled reaching relies on a V2a propriospinal internal copy circuit. *Nature*, 508(7496), 357-363.
5. Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, 37(4-5), 407-419.
6. Balleine, B. W., & O'doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, 35(1), 48-69.
7. Barabási, D. L., Ferreira Castro, A., & Engert, F. (2025). Three systems of circuit formation: assembly, updating and tuning. *Nature Reviews Neuroscience*, 1-12.
8. Bavelier, D., Levi, D. M., Li, R. W., Dan, Y., & Hensch, T. K. (2010). Removing brakes on adult brain plasticity: from molecular to behavioral interventions. *Journal of Neuroscience*, 30(45), 14964-14971.
9. Becker, G. S. (1974). A theory of social interactions. *Journal of political economy*, 82(6), 1063-1093.
10. Bitterman, M.E., Menzel, R., Fietz, A., Schader, S. (1983). Classical conditioning of proboscis-extension in honey bees. *J. Comp. Physiol. Psychol.* 97, 107-119.
11. Brooks, R. A. (1991). Intelligence without representation. *Artificial intelligence*, 47(1-3), 139-159.
12. Changizi, M. A. (2003). Relationship between number of muscles, behavioral repertoire size, and encephalization in mammals. *Journal of Theoretical Biology*, 220(2), 157-168.
13. Chen, W. V., & Maniatis, T. (2013). Clustered protocadherins. *Development*, 140(16), 3297-3302.
14. Chowdhury, M. A. R., An, J., & Jeong, S. (2023). The pleiotropic face of CREB family transcription factors. *Molecules and cells*, 46(7), 399-413.
15. Corbit, L. H., & Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *Journal of Neuroscience*, 25(4), 962-970.
16. Dawkins, R. (1982). *The Extended Phenotype*. Oxford: W. H. Freeman.
17. De Boer, S. F., & Koolhaas, J. M. (2003). Defensive burying in rodents: ethology, neurobiology and psychopharmacology. *European journal of pharmacology*, 463(1-3), 145-161.
18. Deacon, T. W. (1998). *The symbolic species: The co-evolution of language and the brain*. WW Norton & Company.
19. Dennett, D. C. (1984) *Elbow Room*. M. I. T. Press.
20. Dennett, D.C. (1995). *Darwin's Dangerous Idea: Evolution and the Meaning of Life*. New York: Simon and Schuster.
21. Dezfouli, A., & Balleine, B. W. (2012). Habits, action sequences and reinforcement learning. *European Journal of Neuroscience*, 35(7), 1036-1051.

22. Dezfouli, A., & Balleine, B. W. (2013). Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized. *PLoS computational biology*, 9(12), e1003364.
23. Dong, H., Wu, L., Wang, Z., Qu, H., Xie, J., Liu, Y., ... & Jin, Y. (2025). The modern expansion of Dscam1 isoform diversity in *Drosophila* is linked to fitness and immunity. *PLoS Biology*, 23(9), e3003383.
24. Dutrow, E. V., Serpell, J. A., & Ostrander, E. A. (2022). Domestic dog lineages reveal genetic drivers of behavioral diversification. *Cell*, 185(25), 4737-4755.
25. Eaton, R.L. (1968). The predatory sequence, with emphasis on killing behavior and its ontogeny, in the cheetah (*Acinonyx jubatus* Schreber). *Zeitschrift für Tierpsychologie*, 27, 492-504.
26. Eigen, M., & Schuster, P. (1977). The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. *Naturwissenschaften*, 64, 541-565.
27. Eigen, M., & Schuster, P. (1978). The hypercycle: A principle of natural self-organisation part B: the abstract hypercycle. *Naturwissenschaften*, 65, 7-41.
28. Epstein, E. R. (1987). The spontaneous interconnection of four repertoires of behaviour in a pigeon (*Columba livia*). *Journal of Comparative Psychology*, 2, 197-201.
29. Flash, T., & Hochner, B. (2005). Motor primitives in vertebrates and invertebrates. *Current opinion in neurobiology*, 15(6), 660-666.
30. Gahnstrom, C. J., & Spiers, H. J. (2020). Striatal and hippocampal contributions to flexible navigation in rats and humans. *Brain and Neuroscience Advances*, 4, 2398212820979772.
31. Giszter, S. F. (2015). Motor primitives—new data and future questions. *Current opinion in neurobiology*, 33, 156-165.
32. Gould, S. J. (2003). *The Hedgehog, The Fox, and the Magister's Pox*. Three Rivers Press.
33. Gould, J. L., & Gould, C.G. (2007). *Animal Architects: Building and the Evolution of Intelligence*. Basic books.
34. Gray, J.A., & McNaughton, N. (2000). *The Neuropsychology of Anxiety (2<sup>nd</sup> ed.)*. Oxford: Oxford University Press.
35. Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.*, 31(1), 359-387.
36. Haig, D. (2001). *Genomic Imprinting and Kinship*. New Brunswick: Rutgers University Press.
37. Hazlett, B.A. (2007). Conditioned reinforcement in the crayfish *Ocrotectes rusticus*. *Behaviour*, 144, 847-859.
38. Hinton, G.E., & Nowlan, S.J. (1987). How learning can guide evolution. *Complex Systems*, 1, 495-502.
39. Hernández, D. G., Rivera, C., Cande, J., Zhou, B., Stern, D. L., & Berman, G. J. (2021). A framework for studying behavioral evolution by reconstructing ancestral repertoires. *Elife*, 10, e61806.
40. Hull, C. (1943). *Principles of Behavior*. New York: Appleton-Century-Crofts.
41. Jeong, H., Ostrander, E. A., & Kim, J. (2025). Genomic evidence for behavioral adaptation of herding dogs. *Science Advances*, 11(18), eadp4591.
42. Jerison, H. (2012). *Evolution of the brain and intelligence*. Elsevier.
43. Jerome, J., Frantino, E.P., & Sturmey, P. (2007). Backward chaining on the acquisition of internet skills in adults with developmental disabilities. *Journal of Applied Behavior Analysis*, 40, 185-189.
44. Jin, Y., & Li, H. (2019). Revisiting Dscam diversity: lessons from clustered protocadherins. *Cellular and Molecular Life Sciences*, 76(4), 667-680.
45. Kiehn, O. (2016). Decoding the organization of spinal circuits that control locomotion. *Nature Reviews Neuroscience*, 17(4), 224-238.

46. Kremer, M. (1993). The O-ring theory of economic development. *The quarterly journal of economics*, 108(3), 551-575.
47. Lubow, R.E. (1989). *Latent inhibition and conditioned attention theory*. Cambridge: Cambridge University Press.
48. MacPhail, E.M. (1982) *Brain and Intelligence in Vertebrates*, Clarendon Press.
49. Mattfeld, A. T., & Stark, C. E. (2015). Functional contributions and interactions between the human hippocampus and subregions of the striatum during arbitrary associative learning and memory. *Hippocampus*, 25(8), 900-911.
50. Maye, A., Hsieh, C, Sugihara, G., & Brembs, B. (2007). Order in spontaneous behavior. *PLoS one*, 2, e443.
51. Maynard Smith, J. (1987). Natural selection: when learning guides evolution. *Nature*, 329, 761-762.
52. McConnell, P. B. (1990). Acoustic structure and receiver response in domestic dogs, *Canis familiaris*. *Animal Behaviour*, 39(5), 897-904.
53. McDowell, K. A., Hutchinson, A. N., Wong-Goodrich, S. J., Presby, M. M., Su, D., Rodriguiz, R. M., ... & West, A. E. (2010). Reduced cortical BDNF expression and aberrant memory in Carf knock-out mice. *Journal of Neuroscience*, 30(22), 7453-7465.
54. McNaughton, N. (1989). *Biology and emotion*. Cambridge: Cambridge University Press.
55. Ostlund, S. B., Winterbauer, N. E., & Balleine, B. W. (2009). Evidence of action sequence chunking in goal-directed instrumental conditioning and its dependence on the dorsomedial prefrontal cortex. *Journal of Neuroscience*, 29(25), 8280-8287.
56. Owings, D. H., & Coss, R. G. (1977). Snake mobbing by California ground squirrels: adaptive variation and ontogeny. *Behaviour*, 50-69.
57. Pfeifer, R., & Bongard, J. (2006). *How the body shapes the way we think: a new view of intelligence*. MIT press.
58. Pinker, S. (2002). *The Blank Slate: The Modern Denial of Human Nature*. London: Penguin.
59. Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it?. *Cognition*, 95(2), 201-236.
60. Poldrack, R. A., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., & Knowlton, B. J. (2005). The neural correlates of motor skill automaticity. *Journal of Neuroscience*, 25(22), 5356-5364.
61. Prescott, T. J., Redgrave, P. and Gurney, K. (1999) Layered control architectures in robots and vertebrates. *Adaptive Behavior*, 7 (1). pp. 99-127.
62. Punzo, F. (2005). The effect of encounter experience on hunting behavior in the spider wasp, *Pepsis cerberus* Lucas (Hymenoptera; Pompilidae). *Texas Journal of Science*, 57, 165-174.
63. Reznikova, Z., Panteleeva, S., & Danzanov, Z. (2012). A new method for evaluating the complexity of animal behavioral patterns based on the notion of Kolmogorov complexity, with ants' hunting behavior as an example. *Neurocomputing*, 84, 58-64.
64. Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3, 135-156.
65. Roy, D. (2017). Myths about memes. *Journal of Bioeconomics*, 19(3), 281-305.
66. Rundus, A. S., Owings, D. H., Joshi, S. S., Chinn, E., & Giannini, N. (2007). Ground squirrels use an infrared signal to deter rattlesnake predation. *Proceedings of the National Academy of Sciences*, 104(36), 14372-14376.

67. Sahley, C., Rudy, J.W., Gelperin, A. (1981). An analysis of associative learning in a terrestrial mollusc. I. Higher-order conditioning, blocking, and a transient US pre-exposure effect. *J. Comp. Physiol.* 144, 1-8.
68. Sakamoto, K., Karelina, K., & Obrietan, K. (2011). CREB: a multifaceted regulator of neuronal plasticity and protection. *Journal of neurochemistry*, 116(1), 1-9.
69. Shen, W., Liu, H. H., Schiapparelli, L., McClatchy, D., He, H. Y., Yates, J. R., & Cline, H. T. (2014). Acute synthesis of CPEB is required for plasticity of visual avoidance behavior in *Xenopus*. *Cell reports*, 6(4), 737-747.
70. Shine, J. M., & Shine, R. (2014). Delegation to automaticity: the driving force for cognitive evolution?. *Frontiers in Neuroscience*, 8, 90.
71. Skinner, B.F. (1938). *The Behaviour of Organisms*. New York: Appleton-Century-Crofts.
72. Sperry, R. W. (1963). Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proceedings of the National Academy of Sciences*, 50(4), 703-710.
73. Stagkourakis, S., Spigolon, G., Marks, M., Feyder, M., Kim, J., Perona, P., ... & Anderson, D. J. (2023). Anatomically distributed neural representations of instincts in the hypothalamus. *bioRxiv*.
74. Stephens, D. W. (1991). Change, regularity, and value in the evolution of animal learning. *Behavioral Ecology*, 2(1), 77-89.
75. Stephenson-Jones, M., Samuelsson, E., Ericsson, J., Robertson, B., & Grillner, S. (2011). Evolutionary conservation of the basal ganglia as a common vertebrate mechanism for action selection. *Current biology*, 21(13), 1081-1091.
76. Styfhals, R., Seuntjens, E., Simakov, O., Sanges, R., & Fiorito, G. (2019). In silico identification and expression of protocadherin gene family in *Octopus vulgaris*. *Frontiers in physiology*, 9, 1905.
77. Sumbre, G., Gutfreund, Y., Fiorito, G., Flash, T., & Hochner, B. (2001). Control of octopus arm extension by a peripheral motor program. *Science*, 293(5536), 1845-1848.
78. Thorpe, T.W. (1956). *Learning and Instinct in Animals*. Cambridge: Harvard University Press.
79. Tinbergen, N. (1951). *The Study of Instinct*. London: Oxford University Press.
80. Van Iersel, J. J. A., & Tinbergen, N. (1948). " Displacement Reactions" in the Three-Spined Stickleback. *Behaviour*, 1(1), 56-63.
81. Wagner, G. P., & Altenberg, L. (1996). Perspective: complex adaptations and the evolution of evolvability. *Evolution*, 50(3), 967-976.
82. Weir, A. A., Chappell, J., & Kacelnik, A. (2002). Shaping of hooks in New Caledonian crows. *Science*, 297(5583), 981-981.
83. Werkhoven, Z., Bravin, A., Skutt-Kakaria, K., Reimers, P., Pallares, L. F., Ayroles, J., & De Bivort, B. L. (2021). The structure of behavioral variation within a genotype. *Elife*, 10, e64988.
84. Winterbauer, N. E., & Bouton, M. E. (2010). Mechanisms of resurgence of an extinguished instrumental behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, 36(3), 343.
85. Wu, Q., & Maniatis, T. (1999). A striking organization of a large family of human neural cadherin-like cell adhesion genes. *Cell*, 97(6), 779-790.
86. Yang, Q., Logan, D., & Giszter, S. F. (2019). Motor primitives are determined in early development and are then robustly conserved into adulthood. *Proceedings of the National Academy of Sciences*, 116(24), 12025-12034.

87. Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature reviews neuroscience*, 7(6), 464-476.
88. Yue, Y., Meng, Y., Ma, H., Hou, S., Cao, G., Hong, W., ... & Jin, Y. (2016). A large family of Dscam genes with tandemly arrayed 5' cassettes in Chelicerata. *Nature communications*, 7(1), 11252.

## Appendices

### Appendix A: The Baldwin Buffer against Error Catastrophe

Here we outline the connection between behavioural hypercycles and Eigen's original concept of the hypercycle as a solution to error catastrophe. The logic is that learned and genetically encoded behavioural modules interact as quasi-species: complementarity and learning repair increase the effective fidelity with which complex behaviour is maintained. This multiplicative nature is precisely what defines the Hypercycle: it is a chain where the weakest link determines the survival value of the whole (reminiscent of the O-ring production function cited in the main text).

Consider a behavioural sequence decomposed into  $m$  modules (steps) that must all be functionally present for the behaviour to deliver its full fitness payoff  $B$ . Let  $s \in [0,1]$  be the per-module probability that genetic development produces a working module (genetic fidelity) and let  $L \in [0,1]$  be the probability that a missing or faulty module is reconstructed during the lifetime by learning (a lifetime "repair" probability; could be a function of investment in learning). After inheritance and learning, the probability that any single module is functionally available is

$$\begin{aligned} q &= s + (1 - s)L \\ &= L + s(1 - L) \end{aligned}$$

If modules interact complementarily so that the whole sequence only confers fitness when every module works, the probability that an individual acquires the complete functioning sequence is  $q^m$ . Normalizing baseline fitness to 1, a minimal maintenance condition for the behaviour to be sustained by selection is therefore

$$Bq^m \gtrsim 1$$

or at threshold,

$$B(L + s(1 - L))^m = 1$$

Solving this equality for the minimal per-module genetic fidelity required to maintain a sequence of length  $m$  gives

$$s_{\min} = \frac{B^{-\frac{1}{m}} - L}{1 - L}, \quad \left( B^{-\frac{1}{m}} \geq L \right)$$

provided  $B^{-\frac{1}{m}} \geq L$ , ensuring  $s_{\min}$  lies within  $[0,1]$ . Equivalently, for fixed  $s$  and  $L$  the maximal number of modules maintainable is

$$m_{\max} = \frac{\ln\left(\frac{1}{B}\right)}{\ln(L + s(1 - L))}$$

provided the right-hand side is positive or that  $0 < L + s(1 - L) < 1$ .

These expressions make the central mechanisms explicit. When  $L = 0$  (no lifetime repair), we recover the familiar Eigen-type dependence  $s_{\min} = B^{-\frac{1}{m}}$ : longer sequences require exponentially higher per-module fidelity and are prone to "error catastrophe." When  $L > 0$ , the required genetic fidelity falls sharply — even modest learning (small  $L$ ) can greatly increase  $m_{\max}$

because learning substitutes for otherwise exact genetic specification. This is the formal statement that learning + modularity can rescue complexity from the error regime.

### Appendix B: The Rotten Kid as a Quasi-Species

The purpose of this appendix is not to reinterpret Becker biologically, but to demonstrate that the formal closure condition underlying the Rotten Kid result shares the same structure as catalytic closure in a hypercycle. To do this, we sketch a formal framework that maps Eigen's hypercycle, Roy's behavioural chain, and Becker's Rotten-Kid mechanism into a single model. We do not introduce a new economic mechanism here. Rather, we show that Becker's (1974, 1981) Rotten Kid theorem is formally isomorphic to the catalytic closure condition in Eigen's hypercycle. Verbally, in Eigen's hypercycle, each species enhances the replication of the next, and stability arises because catalytic benefits ultimately feed back to the contributor. Becker's Rotten Kid mechanism implements the same closure through transfers rather than chemistry: individual effort raises aggregate output, and transfers tie private payoff to that aggregate. The algebra below shows that these two mechanisms are analogous in structure.

There are  $n$  agents  $i = 1, \dots, n$  arranged in a directed cycle (index arithmetic mod  $n$ ). Each agent chooses an effort  $e_i \geq 0$ . Effort has private cost  $C(e_i)$  with  $C'(e) > 0, C''(e) > 0$ . Each agent's output (or "product") is produced partly by their own effort and partly by catalytic support from the previous agent in the cycle (hypercycle-style). We give this a simple linear specification

$$y_i = ae_i + be_{i-1},$$

where  $a \geq 0$  is the direct productivity of own effort and  $b \geq 0$  is the catalytic effect provided by the predecessor. If  $b > 0$  we have the chain interaction, while if  $b = 0$  agents are independent.

Let the total family/product pool be

$$Y = \sum_{i=1}^n y_i$$

Note that with the linear form  $Y = (a + b) \sum_i e_i$ , a benevolent parent (or allocator) observes the aggregate  $Y$  and commits to a transfer rule  $\phi = (\phi_1(Y), \dots, \phi_n(Y))$  with

$$\sum_i \phi_i(Y) = Y,$$

there is no waste and no other choices are possible. Consider the important canonical case of equal sharing,

$$\phi_i(Y) = \frac{Y}{n} \text{ for all } i$$

This corresponds to Becker's simple example where each child gets an equal share of family output (the analogue of aggregate fitness in the hypercycle). An agent  $i$  chooses  $e_i$  to maximize their own utility:

$$U_i(e_i; e_{-i}) = \phi_i(Y) - C(e_i)$$

Using the chain rule, first-order condition (FOC) becomes

$$\frac{d\phi_i}{dY} \cdot \frac{dY}{de_i} - C'(e_i) = 0.$$

For our linear production,  $\frac{dY}{de_i} = (a + b)$ , each unit of  $e_i$  increases own output by  $a$  and the next agent's output by  $b$ , but summing around the cycle gives  $\frac{dY}{de_i}(a+b)$  per unit of  $e_i$ . So

$$\frac{d\phi_i}{dY}(a + b) = C'(e_i)$$

When redistribution "closes the loop", and if  $\frac{d\phi_i}{dY} > 0$  (transfers increase with aggregate output), then each agent internalizes a *positive marginal return* from increasing the family's aggregate  $Y$ . With equal sharing  $\phi_i(Y) = \frac{Y}{n}$ , we have  $\frac{d\phi_i}{dY} = \frac{1}{n}$  and so the FOC becomes:

$$\frac{a + b}{n} = C'(e_i),$$

which is identical for all  $i$ . Thus, each agent picks the same effort level that equates marginal cost to  $\frac{a+b}{n}$  (by symmetry of preferences and transfers, all agents choose identical effort in equilibrium). Crucially, the agent's marginal benefit in this regime contains the catalytic externality  $b$  (the benefit their effort gives to the next agent) – that externality is internalized, albeit diluted by  $\frac{1}{n}$  (which implies the incentive to redistribute shrinks as family size grows, suggesting a limit for the effective size of the "hypercycle" that can emerge). This is Becker's intuition: because transfers are based on aggregate  $Y$ , even a selfish child chooses effort that increases total output.

This can be compared to no-redistribution (selfish, no father). If an agent keeps only their own output, utility might be

$$\begin{aligned} U_i &= y_i - C(e_i) \\ &= ae_i + be_{i-1} - C(e_i) \end{aligned}$$

Treating  $e_{i-1}$  as fixed, the FOC is:

$$a = C'(e_i)$$

So, the agent ignores the  $b$ -externality. With redistribution, the marginal return is  $\frac{a+b}{n}$  instead of  $a$  which makes two differences: (1) redistribution can *raise* or *lower* marginal incentives depending on parameter values and  $n$ , and (2) redistribution allows private choices to take account of cross-agent catalytic benefits  $b$  that would otherwise be external.

In short, For the Rotten-Kid result to hold, two ingredients are crucial and these conditions define the analogue of 'hypercycle closure' in behavioural terms. First, the parent's transfers must be credible and depend on aggregate outcomes in a way that gives an agent a positive marginal return from increasing the aggregate. Second, agents must be able to affect  $Y$  sufficiently via their own activities (here  $\frac{dY}{de_i} = a + b > 0$ ). Under these conditions, even the "rotten" child who cares only about their transfer will select the effort that (weakly) improves aggregate  $Y$  because it raises their transfer.

Suppose one agent mutates or otherwise comes to contemplate deviating to a low effort  $e'_i$ . The deviation changes total  $Y$  by  $(a + b)(e'_i - e_i)$  and hence their transfer by  $\phi'_i(Y) \cdot (a + b)(e'_i - e_i)$ . They will deviate only if the private gain from lowering effort (reduced cost) exceeds the loss of transfer:

$$C(e_i) - C(e'_i) < \frac{d\phi_i}{dY} (a + b)(e_i - e'_i)$$

With equal sharing, this becomes

$$C(e_i) - C(e'_i) < \frac{a + b}{n} (e_i - e'_i).$$

If  $C$  is convex, this generally fails for a unilateral large shirk: the left side (saved cost) may be larger or smaller than the transfer loss depending on parameters. But in the marginal sense (small deviations), the First Order Condition ensures no profitable infinitesimal deviation. For global stability we need the transfer responsiveness and cost curvature to be such that downward deviations are unprofitable — otherwise the loop can be destabilised.

We can now map these variables easily onto the hypercycle concept. Agents are equivalent to quasi-species with effort corresponding to replication rate and catalytic spillover corresponding to cross-activation coefficients. The catalytic parameter  $b$  is equivalent to the “product” of quasi-species  $i$  that enhances the reproduction/replication of species  $i + 1$ , and the transfer rule  $\phi$  measures a mechanism that “closes the loop” (in Eigen’s hypercycle closure is molecular/catalytic; here it is behavioural by redistribution). The error catastrophe analogues are clear: In hypercycles, a parasitic replicator that takes resources without reciprocating can collapse the cycle. In families, noncredible transfers or preferential treatment that breaks the link between aggregate improvement and individual payoff can let selfish actions proliferate and collapse collective productivity. Interestingly, this perspective suggests a concrete biological route through which the closure conditions required for the Rotten Kid result may fail. If the mechanisms governing contributions or transfers are influenced by imprinted loci, then maternally and paternally derived genes can favour different allocation rules, creating systematic incentives that deviate from aggregate efficiency. As Haig (2001) argues, many such genes regulate tissues that mediate resource allocation (particularly the placenta and the mammalian brain). Intragenomic conflict phenomena may therefore be open to interpretation as a specialized genomic analogue of parasitism in a hypercycle or free riding in the family model, undermining the alignment between individual actions and collective output.