

Transient Phenomena in Learning and Evolution: Genetic Assimilation and Genetic Redistribution

Janet Wiles

James Watson

School of Information Technology
and Electrical Engineering

The University of Queensland
Queensland 4072

Australia

j.wiles@itee.uq.edu.au

jwatson@itee.uq.edu.au

Bradley Tonkes

The University of New South Wales

btonkes@cse.unsw.edu.au

Terrence Deacon

University of California, Berkeley

twdeacon@aol.com

Abstract Deacon has recently proposed that complexes of genes can be integrated into functional groups as a result of environmental changes that mask and unmask selection pressures. For example, many animals endogenously synthesize ascorbic acid (vitamin C), but anthropoid primates have only a nonfunctional version of the crucial gene for this pathway. It is hypothesized that the loss of functionality occurred in the evolutionary past when a diet rich in vitamin C masked the effect of the gene, and its loss effectively trapped the animals in a fruit-eating lifestyle. As a result, the complex of abilities that support this lifestyle were evolutionarily bound together, forming a multilocus complex. In this study we use evolutionary computation simulations to explore the thesis that masking and unmasking can transfer dependence from one set of genes to many sets, and thereby integrate the whole complex of genes. We used a framework based on Hinton and Nowlan's 1987 simulation of the Baldwin effect. Additional gene complexes and an environmental parameter were added to their basic model, and the fitness function extended. The simulation clearly demonstrates that the genetic redistribution effect can occur *in silico*, showing an initial advantage of endogenously synthesized vitamin C, followed by transfer of the fitness contribution to the complex of genes that together allow the acquisition of vitamin C from the environment. As is well known in the modeling community, the Baldwin effect only occurs in simulations when the population of agents is "poised on the brink" of discovering the genetically specified solution. Similarly, the redistribution effect occurs in simulations under specific initial conditions: too little vitamin C in the environment, and its synthesis it is never fully masked; too much vitamin C, and the abilities required to acquire it are not tightly integrated. The Baldwin effect has been hypothesized as a potential mechanism for developing language-specific adaptations like innate universal grammar and other highly modular capacities. We conclude with a discussion of the relevance of genetic assimilation and genetic redistribution to the evolution of language and other cognitive adaptations.

Keywords

Baldwin effect, genetic assimilation, genetic redistribution, genetic variation, adaptation, evolutionary computation

1 Introduction: From Specific Genes to Multilocus Complexes

Vitamin C dependence in humans has an interesting evolutionary history. Many animals have the ability to endogenously synthesize ascorbic acid (vitamin C). The crucial gene in this synthetic

pathway is for an enzyme (LGO) that catalyzes the last stage of synthesis of vitamin C. Anthropoid primates (monkeys, apes, and humans) don't synthesize their own vitamin C, and the question that arises is why not, when it seems such a direct way to ensure adequate supplies of an essential vitamin. It turns out that primates including humans do indeed have the remains of such a gene that has accumulated irreparable mutational damage and is no longer expressed (called a *pseudogene*). It was identified by using a probe gene from the rat [14]. The loss of functionality appears to have happened about 43 million years ago, which is the time when some primates became diurnal, with consequent changes in lifestyle, including a diet containing significant amounts of fruit. A plausible explanation (or just-so story) is that the increased fruit in the animals' diets provided ample regularly available vitamin C, reducing the selection pressure to maintain the function of the vitamin C producing gene. Once the gene was corrupted in this lineage, the animals were effectively addicted to fruit to obtain their vitamin C, and trapped in the fruit-eating lifestyle (*obligate frugivores*). At this point, all the abilities such as color vision, tooth structure, and taste preferences that had incidentally supported the acquisition of vitamin C through diet were subject to a shared and much stronger selection pressure, since the combination of abilities that supported the acquisition of sufficient fruit had become a core component of survival. The original direct selection pressure on the ability to synthesize vitamin C had been replaced by an indirect selection pressure distributed over a set of interrelated abilities to acquire it exogenously. We use the term *genetic redistribution* to refer to the process by which a dependence on an essential nutrient is transferred from the genes in the original pathway for synthesizing it to a complex of genes supporting the ability to acquire it exogenously.

In recent work, Deacon [6] hypothesized that this kind of genetic redistribution is a major force in evolution, applying more generally to the transfer of dependence on a nutrient or behavior from a specific gene to a set of genes that underlie a range of abilities. The process integrates what was initially a diverse set of potentially unrelated skills (such as color vision, tooth structure, and taste preferences, in the vitamin C case) into an adaptive complex that spans multiple loci, supports multiple phenotypic features, and is stable in its own right.

Our goal in this article is to model the genetic redistribution process. We begin by reviewing the relationship between evolution and phenotypic plasticity, in particular the Baldwin effect and the computational studies that have been used to demonstrate it. We then extend those simulations to study the genetic redistribution effect and explore the conditions under which it occurs.

2 Genetic Assimilation and Simulations of the Baldwin Effect

There is a universally held view among biologists that genotypes constitute inherited information that in conjunction with an appropriate environment enables the development of an organism. Changes in phenotype due to use or disuse are not inherited (with the possible exception of the immune system and rare cases of reverse transcription by viral agents), and it is generally agreed that complex behaviors acquired during an individual's lifetime are not incorporated into the genome of its immediate offspring. Thus, the muscles of blacksmiths are not inherited, only the potential to develop them. Similarly, the broken legs of skiers are not inherited by their children, although the potential for broken legs is inherited. The Lamarckian view that giraffe's necks grow longer through reaching into trees and that offspring inherit their parents' longer necks has long been discredited.

In traditional evolutionary theory, to a first approximation, selection acts on phenotypes, genotypes are changed only through mutation and recombination, and mutations are not directed towards the design of any particular organ. However, over many generations, it is clear that complex organs do evolve and are appropriately adapted to the environment of an organism, such as the wide variety of eyes tailored to the different light levels and color spectra for varying habitats. Appropriate variation is essential for the evolution of complex adaptations so that selection has material to work on. In this section we review how increasing and decreasing the amount of variation available in the phenotype can play a role in adaptation.

Waddington used the term “genetic assimilation” to describe a process in which phenotypic-level behavior influences genetic specification (for a review, see [7]). In this process, behaviors that were learned or acquired in the remote past become genetically fixed—not in the subsequent generation, but in the remote descendants. The process sounds Lamarckian, but a Darwinian explanation exists. Over evolutionary time, natural selection acts on phenotypes and under some conditions will act to increase variation (or plasticity—which in this context includes both physical and behavioral plasticity and learning). Under other circumstances natural selection will act to decrease variation.

1. Increased variation or plasticity: If a population is not completely optimized for its current environment, behavioral variation may initially increase the range of behaviors open to individuals or the timing of their development. Behaviors that are initially at the extreme range of a population’s phenotypic plasticity may provide disproportionate benefit to those who can discover and utilize them during their lifetime. Such plasticity is then selected for, until the entire population is composed of individuals with such plasticity.
2. Decreased variation or plasticity: If a population is composed of individuals with high plasticity who can thereby acquire and utilize favorable behaviors, but there is a cost associated with the plasticity, then selection will favor individuals who are better optimized to their environment with less variation. Small mutations that provide a head start for the phenotype towards good solutions will be favored. The head start may be accomplished in a multitude of ways—they all come under the optimization term *bias*. Over time, biases may become so strong that virtually no environmental stimulus is needed, and the behavior can be considered genetically specified. The mechanism that underlies genetic assimilation is thus differential selection of individuals that either need less learning or learn faster.

The idea that plasticity or variation can play a functional role in evolutionary search using purely Darwinian mechanisms was proposed in 1896 by Baldwin [2] and Morgan [13]. Their original motivation was to understand how phenotypic plasticity or learning could guide evolution without resorting to Lamarckianism. This interaction between learning and evolution has come to be known as the Baldwin effect.

From an evolutionary search perspective, phenotypic plasticity is credited with the ability to explore phenotypic space considerably faster and cheaper than genetic search allows. Genetic assimilation provides a Darwinian mechanism that enables the outcomes of phenotypic search to be incorporated into the genome over evolutionary time scales.

A computational model of the Baldwin effect was first demonstrated by Hinton and Nowlan [10]. They designed a simple simulation in which learning could guide an evolutionary algorithm to solve a computational version of a needle-in-haystack task. A population of agents were initialized with random genomes that included alleles for both innate and plastic behavior. Agents were selected for their ability to find a single high-fitness phenotype during their “lifetimes.” Without phenotypic plasticity, the search task is exponential in the number of genes. Populations that are large enough for one or a few individuals to solve the task genetically lose the solution through crossover before it can become genetically fixed. With plasticity, populations reliably find the high-fitness phenotype and genetically assimilate it.

Since Hinton and Nowlan’s original simulation, a generation of computational modelers have explored multiple facets of the Baldwin effect [1, 3, 5, 8, 9, 11, 17, 19, 20]. It has been proposed as a potential mechanism for the evolution of language-specific adaptations [16], and the interplay between learning and evolution is also relevant to embodied cognition and evolutionary robotics [15]. Complex interactions occur between the parameters in the simulations. Of particular importance to this study is that finding and assimilating a solution to an evolutionary search task occurs only under quite stringent circumstances:

- In Hinton and Nowlan’s original model [10], the population size times the space that an agent can search in its lifetime was approximately equal to the size of the entire phenotypic

search space. This balance ensured that in the initial random population, at most a few individuals would find the needle phenotype. Thus, the population was poised on the brink of discovering the solution.

- Populations rarely eliminated learning completely. The fitness function used by Hinton and Nowlan yields very little selection pressure for the complete assimilation of the needle-in-a-haystack task. Residual learning occurs when genetic drift combines with hitchhiking genes from the initial founders and some loci rapidly converge on single alleles (called “homozygous” alleles) [9, 20].
- The selection algorithm (e.g., fitness proportional or tournament selection) makes a difference to the amount of residual learning and the speed of assimilation [21].
- The Baldwin effect has been demonstrated with a variety of models, including neural networks [5, 19]. The rate of environmental change and the complexity of the landscape affect the benefits of learning in a rugged landscape [18].

The interactions between parameters in the Baldwin effect can be summarized in terms of two components, the benefits and costs of learning [11]. In the early stages of a simulation, phenotypic plasticity is a cost-effective way to search the local fitness landscape, compared to genetic search. At later stages, the costs of learning outweigh the benefits if all individuals can learn. However, residual learning is only eliminated if the learning costs remain high.

In summary, the genetic assimilation of acquired abilities can be explained in Darwinian terms, changing frequencies of alleles over a population of learners over generations. Exploration with computational models indicates that learning systems need not produce both stages of the Baldwin effect. In some cases learning does not increase the speed of finding a solution, nor does it result in genetic assimilation. Both stages of the Baldwin effect only occur when learning changes the selection of parents for the subsequent generation, first as an advantage, and then as a cost.

The conclusion we draw from these studies is that the rise and fall of learning in a population is a specific transient dynamic in the evolutionary process, not a ubiquitous feature of phenotypic plasticity.

3 Modeling Genetic Redistribution

Viewing the Baldwin effect as a transient phenomenon in evolutionary dynamics raises the possibility that other such phenomena may occur. The vitamin C story outlined above is a plausible example. Deacon’s [6] thesis that complexes of genes can be integrated into functional groups as a result of environmental changes that mask and unmask selection pressures lends itself to exploration using modeling techniques analogous to the Baldwin effect simulations.

When endogenous synthesis of vitamin C is masked by its presence in the diet, the change in selection pressure can be viewed as causing a “reverse Baldwin effect” in that abilities specified directly in the genome may become masked (by internal or external sources, including flexible behavioral abilities), so that over time their genetic specification is lost. When this occurs, the individual becomes dependent on the source that provided the masking effect, and any phenotypic capacities that support this masking (e.g., by providing an externally redundant nutrient) become increasingly elaborated and integrated through positive selection pressure. Deacon describes the event that results in the loss of direct genetic specification of an ability as *masking* selection for the original genes, and the transfer of dependence for the ability to the masking source as *unmasking* of selection on the new complex of genes.

In this study we use evolutionary computation simulations to explore the thesis that masking and unmasking can transfer dependence from one set of genes to many sets, and thereby integrate the whole complex of genes.

The first aim of the simulation was to demonstrate the genetic redistribution effect *in silico*, showing an initial advantage of endogenously synthesized vitamin C, followed by transfer of the benefit to the complex of genes that together allow the acquisition of vitamin C from the environment. The second aim of the model was to explore the conditions under which the phenomenon occurs and its sensitivity to contributing factors and parameters.

3.1 Details of the Model Design

The design is an evolutionary computation (EC) model of genetic change over generations. The architecture is based on Hinton and Nowlan’s [10] simulation. Additional gene complexes and an environmental parameter were added to their basic model, and the fitness function extended. The elegance of the architecture is that it uses their framework but shows a different effect by changing the environmental context.

Hinton and Nowlan used a single environment, which was constant throughout the simulation. In the current study, the environment is modeled as a parameter that specifies the maximum amount of vitamin C available in an agent’s environment. The environment has three stages, starting at a very low base, then rising to more than required for daily use, then dropping to half the daily limit.

To benefit from the environment, the agent must have the full set of abilities to acquire vitamin C from its food. If even one of those abilities was missing, then no exogenous vitamin C was acquired. The total vitamin C for an agent is calculated as the sum of the endogenous and exogenous sources, up to a maximum value. Any vitamin C beyond the maximum is discarded.

Following Hinton and Nowlan, in the current study each ability was modeled using a chromosome of 20 genes per ability, and the fitness contribution was based on a needle-in-a-haystack task (see Figure 1). To distinguish between the genotype and phenotype, the expressed value of a gene will be referred to as a *phene*. Each gene had three alleles: correct (1), incorrect (0), and learnable (?). The correct and incorrect alleles were fixed, and could not change during an agent’s lifetime. The learnable alleles were reset randomly to correct or incorrect phenes each day of an agent’s life. Each phene thus had two alleles, correct and incorrect. The agent was credited with a fitness contribution for an ability only when all its phenes were correct (no distinction was made between phenes that were innately correct and those that were corrected through guessing). If the ability contained one or more incorrect (0) alleles, the fitness contribution of each ability was zero; if all the alleles were innately correct, the fitness contribution was maximum; otherwise, the fitness contribution reflected a penalty for the expected number of guesses taken to find the correct phenotype (which is 2^x , where x is the number of learnable (?) alleles).

In the vitamin C model, setting the first 20 genes to 1 corresponds to the ability (A_0) to synthesize vitamin C endogenously. The additional abilities needed to acquire vitamin C from the environment were modeled as k additional gene complexes (A_1, A_2, \dots, A_k). Each additional ability was also modeled as the coordinated action of 20 genes (resulting in $N = 20(1+k)$ genes in the chromosome). Abilities $A_1 - A_k$ were initially independent and individually advantageous (e.g, color vision, tooth structure, taste preference). Together they enabled the agent to acquire and utilize vitamin C from its environment. For all the simulations in this study, k was set to 3.



Figure 1. Genotype for an agent. Each agent’s chromosome is a vector of 80 genes; each gene has allele 0, 1, or ?. The phenotype for an agent is a vector of 80 phenes; each phene can be 0 or 1, either directly specified from the corresponding gene or, where the gene is learnable (?), randomly assigned. See text for details.

The fitness function F contained terms for the total vitamin C and the sum of the additional abilities:

$$F = 1 + (N - 1)(c + a)$$

where c is total vitamin C from both endogenous and environmentally derived sources, with excess vitamin C discarded, and a is the total incidental fitness from the other abilities (see Table 1 for details of the fitness function).

3.2 Methods

A simulation trial consisted of generating a population of 1,000 agents, each with the genome described above (see Figure 1), and all genes initialized to random values with probability 35% for 0's, 25% for 1's, 40% for ?'s. The fitness of all agents was evaluated (see Table 1) and used to select parents for the next generation (roulette wheel or tournament selection [12]). A new generation was created by repeatedly selecting two parents and using single-point crossover to create two offspring. All genes in the new agents had a $1/N$ chance of mutating (using the same probabilities as in the initial distribution).

Each simulation trial was run for 2,000 generations, divided into three stages, depending on the amount of vitamin C in the environment, E . The first stage (0–700 generations) modeled the prehistory of the population, during which all abilities had the opportunity to be found. The amount of vitamin C available in the diet was set to a low value (5% of the usable amount). It was expected that each ability would be found with a time course independent of the other abilities and no particular order would be observed in their discovery. It was considered possible that some trials would suffer from hitchhiking genes and premature convergence, preventing all abilities from being discovered. Such trials were aborted at stage 2. The second stage (700–1,200 generations) was designed to show the effect of masking the synthesis of vitamin C by increasing the vitamin C in the diet to an overabundance (200% of the usable amount). It was expected that mutational damage would accumulate in ability A_0 and consequently the fitness contribution a_0 would decrease. The third stage (1,200–2,000 generations) was designed to show the effect of increasing the selection pressure on the complex of abilities (A_1 – A_3) by reducing the vitamin C in the diet to make it rarer but still available (50% of the usable amount). It was expected that there would be selection pressure to optimize abilities A_1 – A_3 , and that this would be demonstrated by a correlated rise in the fitness contributions r_1 , r_2 , r_3 . It was also expected that the population would have converged at this stage and that the rediscovery of A_0 would be much less probable than at the start of a trial, due to the lack of required variation in the genes.

Two sets of simulations were run, one consisting of 100 trials using fitness proportional selection, and the other consisting of 100 trials with tournament selection. Measurements were taken of the total fitness of the population, F , and the individual contributions to the fitness, a_0 , r_1 , r_2 , and r_3 .

In summary, the first hypothesis of the simulations was that the transition from low to high levels of vitamin C in the environment would mask the contribution of the ability to synthesize vitamin C, as shown by a rapid drop in the fitness contribution a_0 at the start of stage 2. The second hypothesis was that the transition from high to lower levels of vitamin C in the environment would unmask the complex of abilities that contribute to obtaining vitamin C from the environment, as shown by a correlated rise in the fitness contributions r_1 – r_3 at the start of stage 3 (in contrast to uncorrelated rises in r_1 – r_3 in stage 1). The third hypothesis was that the complex of abilities for obtaining vitamin C from the diet would block or reduce the probability of rediscovering the ability to synthesize vitamin C, as shown by the fitness contribution a_0 remaining at zero throughout stage 3.

3.3 Results

Hypotheses 1 and 2 were supported in all complete trials for both simulation methods. Hypothesis 3 was partially supported: Blocking was observed as predicted in all complete trials for fitness

Table 1. Fitness function.

The fitness F of an agent depends on the number of genes in ability A_0 , the total vitamin C (endogenous and acquired), and the other abilities:

$$F = 1 + (N - 1)(c + a)$$

where

N is the genome length of A_0 ;

c is the total vitamin C (endogenous and acquired), given by

$$c = \text{Min}[c_0 + E_{c1}, 1.0];$$

E is the amount of vitamin C available in the environment;

c_0 is the amount of endogenous vitamin C, and c_1 is the ability to acquire vitamin C, given by

$$c_0 = \begin{cases} \text{Max}[0, (M - 2^x)/M] & \text{if } A_0 \text{ has no 0s, and } x \text{ is the number of ?s in } A_0 \\ 0 & \text{otherwise} \end{cases}$$

$$c_1 = \text{Min}[r_1, r_2, r_3]$$

M is the number of guesses;

a is the total contribution of other abilities, given by $a = \lambda \sum r_i^*$;

λ is the abilities constant, which scales the relative contributions of vitamin C and the other abilities;

r_i^* is the fitness contribution from ability A_i , given by

$$r_i = \begin{cases} \text{Max}[0, (M - 2^y)/M] & \text{if } A_i \text{ has no 0s, and } y \text{ is the number of ?s in } A_i \\ 0 & \text{otherwise} \end{cases}$$

$$r_i^* = \text{Min}[r_i, 0.75]$$

The simulation parameters for this example were set as follows: $N = 20$, population size = 1000, and $M = 1000$ following Hinton and Nowlan [11]. The distribution of alleles (initial and mutation) was set to 35% 0's, 25% 1's, 40% ?'s; this distribution was chosen to make the learning task slightly harder than Hinton and Nowlan's. The number of abilities, k , to acquire vitamin C from the environment was chosen as the minimum that could feasibly represent a complex of abilities, $k = 3$. The scaling parameter λ was chosen to provide a balance between the contributions of the different abilities, $\lambda = 30/(N - 1)$. The simulations used tournament selection or fitness proportional selection with mutation and single-point crossover.

proportional selection, but was only observed in a portion of trials using tournament selection. Successful trials (showing all three predictions) showed very similar behavior: The fitness contributions showed the discovery of all four abilities, A_0 – A_3 , during stage 1, with their time courses for discovery uncorrelated, the rapid loss of ability A_0 during stage 2, and a correlated rise in the fitness contributions of A_1 – A_3 at the start of stage 3 and blocking of the rediscovery of A_0 (see Figure 2).

The time course of the successful simulations demonstrates the initial advantage of the endogenously synthesized vitamin C, followed by a transfer of the ability to the complex of genes that mask the effect. The three stages can be summarized:

1. *Prehistory (0–700 generations)*: All four abilities are found.
2. *Masking (700–1,200)*: Ample vitamin C in the environment results in rapid loss of ability A_0 .
3. *Increased selection pressure (1,200–2,000)*: Vitamin C in the environment is reduced, resulting in a coordinated increase in the fitness contributions of A_1 – A_3 .

Stage 1 is the prehistory of the effect, in which the population is evolving all its abilities (see Figure 2a–f, generations 0–700). The ability to synthesize vitamin C, A_0 , is the first ability acquired in this trial (Figure 2c), followed by A_1 to A_3 (Figure 2d–f). In other trials they are found in random order. The important point to note in this stage is that each ability is found independently (in that the generation in which they were discovered is uncorrelated), and hence there is no coordination in the timing of their fitness increases. All abilities make a small contribution to the total fitness, which is why they evolve in the first stage. When all three abilities A_1 – A_3 are found, the population has the potential to utilize vitamin C in the environment, but in this first stage there is virtually no environmental source, and hence no additional benefit beyond the individual contributions to fitness.

Stage 2 is the masking stage and begins when the environment changes to include significant amounts of vitamin C (see Figure 2a–f, generations 700–1,200). During this stage, the dietary vitamin C masks the contribution of endogenous vitamin C. Random mutations gradually accumulate in the genes of A_0 , and the ability to synthesize vitamin C is lost. At this point, the population is addicted to exogenous vitamin C. Abilities A_1 – A_3 , which initially evolved independently of one another, are now bound together. At this stage, their co-dependence is not directly visible from the graphs, but any further evolutionary change must take into account that their combined effect is essential in acquiring vitamin C (modeled in the fitness function as the minimum of the contributions r_1 – r_3). However, the excess vitamin C in the environment enables them to be less than optimal in their acquisition of it. In Figure 2b, at 700 generations, a slight drop in fitness is seen. This average reflects loss of robustness to mutation, with 75% of the population at maximum fitness and 25% with mutations.

Stage 3 reveals that selection pressure on the complex of abilities has been unmasked, and begins when the conditions change again and the amount of vitamin C in the environment is reduced to half the maximum benefit (see Figure 2a–f, generations 1,200–2,000). There is now selection pressure for abilities A_1 – A_3 to optimize their extraction of vitamin C from the environment, and a coordinated increase in their fitness is seen (shown as a coordinated increase in fitness over generations 1,200–1,250 in Figure 2d–f). Blocking of the rediscovery of vitamin C synthesis occurs because the chance of an agent rediscovering the correct genes for A_0 is significantly less than during stage 1. The reason is a much lower level of variation in the alleles of the population, and convergence of some loci to incorrect values.

The successful trials that completed all three stages are taken as demonstration of the genetic redistribution effect. However, one of the features of the genetic algorithm methodology employed is that relatively few trials completed the full 2,000 generations. Using tournament selection, 30 trials found all four abilities within 1,200 generations, and using fitness proportional selection, only 5 trials did so. In hindsight, this result should have been considered as a possibility, if not actually expected, as drift and convergence have long been observed in Baldwin effect simulations [3, 9,

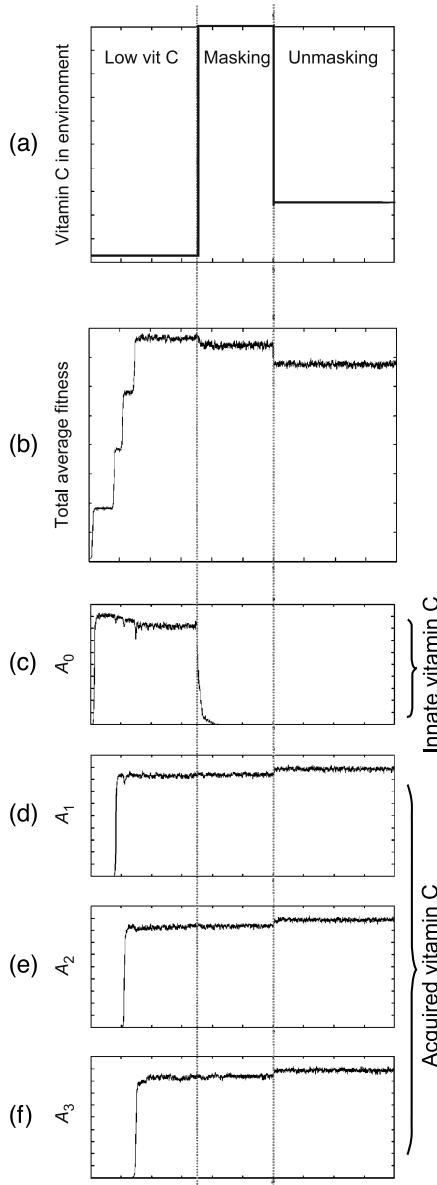


Figure 2. (a) Environmental benefit versus generation. For 700 generations, there is very little vitamin C in the environment. Then the amount rises to 200% of the usable amount. At 1,200 generations it drops back to 50% of the usable amount. (b) Total average fitness F . The population starts with random alleles in the agent’s genomes, and very low fitness. During stage 1, all four abilities are found. At 1,200 generations, the fitness drops markedly, reflecting the reduction of vitamin C in the environment and a reduction of fitness over the population as a whole. (c) Fitness contribution c_0 of ability A_0 . A_0 is rapidly discovered and spreads throughout the population, shown by the rapid rise in c_0 . At generation 700, c_0 plummets, due to a lack of selection pressure to maintain the ability. (d–f) Fitness contributions, $c_1–c_3$ of abilities $A_1–A_3$. $A_1–A_3$ are discovered in the first 700 generations, shown by the rapid rises in their corresponding fitness contributions, and the generations at which the increases occur are uncorrelated. At generation 1,200, they show a correlated rise in fitness, optimizing their extraction of vitamin C from the environment.

10, 20]. The genome for the current simulations is four times the size of the original model (80 genes compared to 20 originally), with a much greater opportunity for loss of diversity in the alleles over the population in one or more loci due to hitchhiking and drift. From previous experience with the Baldwin effect, it seems likely that in trials using fitness proportional selection, hitchhiking and rapid convergence of the population frequently prevent all four abilities being found (95% of trials were aborted), and that in trials using tournament selection, convergence due to prolonged periods of drift and insufficient selection pressure may prevent all abilities being found within the given time frame (70% of trials were aborted). See Wiles et al. [20] for further discussion of hitchhiking and drift in the Baldwin effect using different selection methods.

Of the complete trials, the method of selection did show a difference in the proportion of times that vitamin C was blocked from re-evolving. Using fitness proportional selection, all completed trials blocked the return of \mathcal{A}_0 . Using tournament selection, 43% (13/30) of trials blocked the return of \mathcal{A}_0 . These differences reflect the relative difficulty of finding all four abilities in the different selection methods.

4 Discussion and Conclusions

The simulations clearly demonstrate the effects of masking and unmasking *in silico* and the consistency of the argument for genetic redistribution. In particular, two main aspects of the simulation were robust to a variety of parameters, which were tested in a range of comparison trials:

- Endogenous synthesis of vitamin C (\mathcal{A}_0) was consistently lost due to masking by other abilities and an alternative source in the environment (transition from stage 1 to stage 2). In trials that were given sufficient vitamin C in the environment, loss of \mathcal{A}_0 was complete in all cases. In other trials in which the source of vitamin C in the environment was insufficient, only partial loss of \mathcal{A}_0 was observed, and it was quickly recovered when unmasked.
- There was coordinated increase in the fitness contributions of \mathcal{A}_1 – \mathcal{A}_3 , showing that unmasking can bind together multiple abilities (transition from stage 2 to stage 3). The abilities are bound in the sense that they have lost the ability to evolve independently, and each must maintain its contribution to the acquisition of vitamin C from the diet. The increase in fitness seen at the transition only occurs if the abilities can be optimized for acquiring vitamin C beyond the fitness required in their original function.

These two transition phenomena can be predicted reliably from the amount of selection pressure in each stage of the simulation. Masking is due to a loss of selection pressure when the environmental source of vitamin C enables other abilities to acquire the daily limit of vitamin C. Unmasking by lower levels of vitamin C in the environment increases the selection pressure to acquire as much of the environmental source as possible, and hence increase the fitness contributions of \mathcal{A}_1 – \mathcal{A}_3 .

What can simulations such as Hinton and Nowlan's and ours show? Like Hinton and Nowlan's simulation, the evidence that we present here demonstrates the consistency of the arguments *in silico* and enables observation of the behavior of the model under different environmental conditions and different parameter choices. The model facilitates the study of the conditions under which evolution's arrow causes genetic assimilation, and the conditions that cause specificity to be lost and genetic redistribution to occur.

The Baldwin effect shows how learning can guide evolution. The genetic redistribution effect shows how masking and unmasking by environmental changes can transfer selection from specific genetic abilities to distributed suites of abilities, and binds those abilities into evolutionarily synchronized components. Both can be viewed as transient phenomena in the evolutionary dynamics of evolving complex phenotypes stimulated by different environmental conditions.

The field of complex systems studies the relationships between different levels of description of systems, and tries to characterize how emergent properties at one level arise from their component parts. Conventional explanations don't usually have causality encompassing very different levels of a multilevel system. When phenomena seem to unexpectedly influence an entirely different level, such as acquired abilities influencing the genetics of their distant descendents, Cohen and Stewart [4] call it a *strange loop* and describe the Baldwin effect as one such strange loop. Genetic redistribution can be considered another such strange loop in the understanding of evolution and learning. (Note that this use of the term differs from Hofstadter's use in *Gödel, Escher, Bach*.)

5 Implications for Cognitive Science

It has been hypothesized that the Baldwin effect is a potential mechanism for evolving specific language adaptations for universal grammar and other modular capacities:

The Baldwin effect probably played a large role in the evolution of brains. Contrary to the standard social science assumptions, learning is not some pinnacle of evolution attained only recently by humans. All but the simplest animals learn. . . . If the ability to learn was in place in an early ancestor of the multicellular animals, it could have guided the evolution of nervous systems toward their specialized circuits even when the circuits are so intricate that natural selection could not have found them on its own.

Pinker [16: p. 179]

From a cognitive modeling perspective, both the Baldwin effect and the conditions under which it occurs are important. However, the existence of learning alone does not determine whether a Baldwin effect will occur. Other conditions must also be satisfied. The evolution of specialized circuits will only benefit from the transient dynamic when the population is poised on the brink of discovery and the benefits and costs of learning are appropriately balanced.

Few people would argue whether the ability to learn language is in the genes. Clearly genes play a part. The argument in psycholinguistics is whether there are *specific* grammar genes, or whether there is a distributed learning system that is used for many cognitive abilities, not just learning languages. The psycholinguistic evidence is beyond the scope of this article, but the issue illustrates why the interactions between evolution and learning play a role in cognitive science.

Interestingly, Hinton and Nowlan's model of the Baldwin effect does not preclude the evolution of distributed abilities, as shown in the simulations in this study. Applied to the language argument, our simulations of the genetic redistribution effect support such a distributed evolutionary scenario. Deacon [6] has argued that symbolic communication is a powerful masking agent that should transfer genetically specified abilities to distributed suites of abilities, and bind those abilities into evolutionarily synchronized cognitive components. The simulations in this article demonstrate how such an effect could occur. In our simulations, binding is revealed by unmasking, but unmasking is not a necessary condition for it to occur.

In conclusion, evolutionary computation is not a methodology that can resolve the issue of whether language is based on distributed or specific abilities. What it can show is that the underlying phenomena of genetic assimilation and genetic redistribution are both computationally coherent, and simulations can be used to investigate the conditions under which—in *silico*—the transient dynamics occur.

Acknowledgments

We thank Luis Rocha and three anonymous referees for their detailed comments on the earlier version of this article. This work was supported by grants from the ARC to the first author and a UQ postgraduate award to the second author. The third author was studying at the University of Queensland.

References

1. Ackley, D. H., & Littman, M. L. (1992). Interactions between learning and evolution. In C. G. Langton, C. Taylor, C. D. Farmer, & S. Rasmussen (Eds.), *Artificial Life II* (pp. 487–509). Reading, MA: Addison-Wesley.
2. Baldwin, J. M. (1896, 1996). A new factor in evolution. *American Naturalist*, 30, 441–451. Reproduced in R. K. Belew, & M. Mitchell (Eds.), *Adaptive individuals in evolving populations*. Reading, MA: Addison-Wesley.
3. Belew, R. K. (1990). Evolution, learning and culture: Computational metaphors for adaptive search. *Complex Systems*, 4(1), 11–49.
4. Cohen, J., & Stewart, I. (1994). *The collapse of chaos*. London: Penguin.
5. Chalmers, D. (1990). The evolution of learning: An experiment in genetic connectionism. In D. S. Touretsky, J. L. Elman, T. J. Sejnowski, & G. E. Hinton (Eds.), *Proceedings of the 1990 Connectionist Summer School* (pp. 81–90). San Mateo, CA: Morgan Kaufmann.
6. Deacon, T. W. (2003). Multilevel selection in a complex adaptive system: The problem of language origins. In B. H. Weber, & D. J. Depew (Eds.), *Evolution and learning: The Baldwin effect reconsidered*. Cambridge, MA: MIT Press.
7. Frazzetta, T. H. (1975). *Complex adaptations in evolving populations*. Sunderland, MA: Sinauer.
8. French, R. M., & Messinger, A. (1994). *Genes, phenes and the Baldwin effect: Learning and evolution in a simulated population*. In R. Brooks & P. Maes (Eds.), *Artificial Life IV*. Cambridge, MA: MIT Press.
9. Harvey, I. (1993). The puzzle of the persistent question marks: A case study of genetic drift. In S. Forrest (Ed.), *Proceedings of the Fifth International Conference on Genetic Algorithms*. San Mateo, CA: Morgan Kaufmann.
10. Hinton, G. E., & Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems*, 1, 495–502.
11. Mayley, G. (1996). The evolutionary cost of learning. In P. Maes, M. J. Mataric, J.-A. Meyer, J. Pollack, and S. W. Wilson (Eds.), *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*. (pp. 458–467). Cambridge, MA: MIT Press.
12. Mitchell, M. (1996). *An introduction to genetic algorithms*. Cambridge, MA: MIT Press.
13. Morgan, C. L. (1896). On modification and variation. *Science*, 99, 733–740.
14. Nishikimi, M., Fukuyama, R., Minoshoma, S., Shimizu, N., & Yagi, K. (1994). Cloning and chromosomal mapping of the human nonfunctional gene for L-gulonolactone oxidase, the enzyme for L-ascorbic acid biosynthesis missing in man. *Journal of Biological Chemistry*, 269, 13685–13688.
15. Nolfi, S., & Floreano, D. (2000). *Evolutionary robotics: The biology, intelligence, and technology of self-organizing machines*. Cambridge, MA: MIT Press.
16. Pinker, S. (1997). *How the mind works*. London: Penguin.
17. Turney, P., Whitley, D., & Anderson, R. (Eds.) (1996). *Evolution, learning and instinct: 100 years of the Baldwin effect*. Cambridge, MA: MIT Press.
18. Watson, J., Geard, N. & Wiles, J. (2002). Stability and task complexity: A neural network model of evolution and learning. In R. K. Standish, M. A. Bedau, & H. A. Abbass (Eds.), *Proceedings of the Eighth International Conference on Artificial Life* (pp. 153–156).
19. Watson, J., & Wiles, J. (2002). The rise and fall of learning: A neural network model of the genetic assimilation of acquired traits. In X. Yao (Ed.), *Proceedings of the 2002 Congress on Evolutionary Computation* (pp. 600–605). Piscataway, NJ: IEEE.
20. Wiles, J., Schulz, R., Bolland, S., & Hallinan, J. (2001). Probing the persistent question marks. In L. Spector, E. Goodman, A. Wu, W. B. Langdon, H.-M. Voigt, M. Gen, S. Sen, M. Dorigo, S. Pezeshk, M. Garzon, & E. Burke (Eds.), *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001)* (pp. 710–717). San Mateo, CA: Morgan Kaufmann.
21. Wiles, J., Schulz, R., Bolland, S., Tonkes, B., & Hallinan, J. (2001). Selection procedures for module discovery: Exploring evolutionary algorithms for cognitive science. In J. D. Moore, and K. Stenning, (Eds.), *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 1124–1129). Mahwah, NJ: Erlbaum.