Organic Selection and Social Heredity: The *Original* Baldwin Effect Revisited

Nam Le

Natural Computing Research & Applications Group
University College Dublin, Dublin, Ireland, A94 XF34
namlehai90@gmail.com

Abstract

The so-called "Baldwin Effect" has been studied for years in the fields of Artificial Life, Cognitive Science, and Evolutionary Theory across disciplines. This idea is often conflated with genetic assimilation, and has raised controversy in trans-disciplinary scientific discourse due to the many interpretations it has. This paper revisits the "Baldwin Effect" in Baldwin's original spirit from a joint historical, theoretical and experimental approach. Social Heredity – the inheritance of cultural knowledge via non-genetic means in Baldwin's term – is also taken into account. I shall argue that the Baldwin Effect can occur via social heredity without necessity for genetic assimilation, instead the Baldwin Effect can promote more plasticity to facilitate future intelligence when the fidelity of social heredity is high. Computational experiments are then carried out to support the hypothesis of interest. The role of mind and intelligence in evolution and its implications in an extended synthesis of evolution are briefly discussed.

Introduction

Studying the relationship between evolution and learning is a very important topic in understanding adaptive behaviour demonstrated by both natural and artificial agents. There exists an intriguing idea called the **Baldwin Effect** by Simpson (1953), named after James Mark Baldwin, as an interpretation of *Organic Selection* proposed by Baldwin (1896). Since Simpson, this idea has often been interpreted as how an adaptive behaviour first acquired during lifetime can later be replaced by fixed innate traits due to the cost of individual learning. This interpretation has often conflated the Baldwin Effect with *genetic assimilation* by Waddington (1953). The Baldwin Effect studied in ALife and complex adaptive systems often used the interpretation of Simpson (Hinton and Nowlan (1987), Harvey (1996), Mayley (1997)).

This interpretation, intentionally or unintentionally, has made the Baldwin Effect more restrictive than what Baldwin originally proposed through organic selection. Most studies in this line of thought, including those in ALife, often neglected the presence and importance of **social heredity** – what Baldwin (1896) originally meant by a parallel heredity of social knowledge via non-genetic means. Baldwin's ideas of social heredity and its influence on evolution bear some sort of similarity to what we now call *gene-culture coevolution* or *dual-inheritance theory* (Peter J. Richerson (2005)).

When social heredity comes in, the story would be more interesting as to how the Baldwin Effect occurs. Some of the interesting questions could be asked as: if social heredity is permitted, then if adaptive information can be gained and transmitted easily through social transmission, what would genetic assimilation look like? More curiously, is genetic assimilation necessary to claim the presence of the Baldwin Effect as studied previously? Plausibly, it seems to us that if adaptive behavioural information is encoded into culture, and this information can be handed down easily from generation to generation by some form of social learning, *ontogenetic* learning still plays a role in directing evolution but the assimilation step seems not to be required. After 'the new factor' in Baldwin (1896), Baldwin stressed the importance of social heredity more in his later books (Baldwin (1902), Baldwin (1909)) that I think it is worth a further investigation to understand what he really meant by his *effect*.

The main aim of this paper is to re-discuss the Baldwin Effect in Baldwin's original spirit to clarify what the *effect* would possibly be. My contribution in this paper can be divided into two parts. First, I present and discuss the history of understanding to show why and how Baldwin's original idea, and ideas, may differ from the rich literature studying the Baldwin Effect. I will prove that Baldwin did not restrict his new factor in evolution to the idea of genetic assimilation, instead he believed that social heredity can provide another way to affect evolution, which may promote plasticity to boost the intelligence of an evolving system. Second, a simple computer simulation, combining evolution, learning, and cultural inheritance, is carried out in order to see how this combination affects the underlying evolutionary process, and whether this would-be effect requires a strict requirement of genetic assimilation. The last section briefly presents some future implications of the Baldwin Effect in various avenues, including the present-day interest in extending and expanding Darwinian account of evolution into a new synthesis (Laland et al. (2015)).

The Baldwin Effect

For clarity, I shall use the term *Baldwin-Baldwin Effect* to refer to the Baldwin's original effect, and Baldwin-Simpson Effect as a reference to Simpson's re-interpretation.

I. A Brief History of Understanding

A. The Baldwin-Baldwin Effect:

At the turn of the 20th century, the idea that learning as part of the ontogenetic adaptation can influence, and somehow *direct* an evolutionary process without resorting to Lamarckian inheritance, was proposed independently by at least three independent thinkers: Baldwin (1896) (published in The American Naturalist), Osborn (1896), and Morgan (1896) (both published in Science Magazine). Baldwin (1896) rediscussed and joined his two previous ideas, published in *Mental Development in the Child and the Race* (Baldwin (1895)), on *Organic Selection* (chap. vii) and *Social Heredity* (chap. xii), and called this *"A new factor in evolution"*.

When first appeared, the idea by Baldwin (also Morgan and Osborn) set a new movement in understanding how evolution works, more specifically when it comes to explaining the inheritance of acquired characteristics. Before Baldwin and the like, the French Naturalist Lamarck proposed that characters acquired during the lifetime of the parent are directly passed down onto the offspring. The English Philosopher Herbert Spencer seemed to agree with Larmarckian inheritance when he said "intelligence would allow an animal to acquire complex habits that would later solidify into instincts. But such transformation required Lamarckian inheritance" (Richards (1989)). Darwin himself believed that Lamarckian evolution might play a small role in life, but most Darwinians rejected Lamarckism (Huxley (1942)) based on Weismann et al. (1893).

Baldwin came to light and explained evolution without resorting to Lamarckian style, in which acquired characters are somehow *indirectly* inherited. A new factor in evolution by Baldwin is organic selection, which includes any form of individual adaptation during the lifetime (through Physico-genetic, Neuro-genetic Psycho-genetic) that directs the evolutionary pathway of an evolving species. He stressed the role of psycho-genetic, by which he meant conscious intelligence, that includes any form of ontogenetic learning, such as imitation, pleasure and pain, reasoning. For Baldwin, it is organic selection that can explain how a behaviour that has learned might be becoming *innate*, or *partially innate*, in future generations. If a group of animals migrates into a new environment for which they initially lack congenital adaptations, those **plastic** enough to accommodate themselves through conscious learning will tend to survive, blocking the strong hand of natural selection. This will allow natural selection opportunity to accumulate chance variations that follow the path laid down by the acquired behaviours. Acquired characteristics are immediately heritable implied a loss of phenotypic flexibility. Such inheritance would tend so to bind up the childs nervous substance in fixed form that he [or she] would have less or possibly no plastic substance left to learn with.

Interestingly, Baldwin did insist the importance of what he termed **Social Heredity** – a means of extra-organic transmission from generation to generation through copying, imitation, teaching, or any form of social learning. Baldwin (1896) considered it heredity because of the following reasons: 1) it is a handing down of physical functions; while it is not biological (physical) heredity; 2) it directly influences physical heredity in the way mentioned, i.e., it keeps alive variations, thus sets the direction of ontogenetic adaptation, thereby influencing the direction of the available congenital variations of the next generation. Of course, social heredity is a form of organic selection or ontogenetic adaptation, but it deserves a special name because of its special way of operation and its farther value. It keeps alive a series of functions which either are **not yet, or never do become, congenital at all**.

Fixity or Plasticity

Baldwin (1896) said: "The two ways of securing development in determinate directions – the purely extra-organic way of Social Heredity, and the way by which Organic Selection in general (both by social and by other ontogenetic adaptations) secures the fixing of phylogenetic variations, as described above – seem to run parallel". And more importantly he concluded that in more complex living animals like humans, *"social transmission is an important factor, and the congenital equipment of instincts is actually broken up in order to allow the plasticity which the human being's social learning requires him to have"*.

Later in *Development and Evolution* Baldwin (1902) said "organic selection opens a great sphere for the application of the principle of natural selection among organisms, i.e. selection on the basis of what they do rather than what they are; of the new use they make of their functions rather than of the mere possession of certain congenital characteristics. A premium is set on plasticity and adaptability of function rather than on congenital fixity of structure; and this adaptability reaches its highest levels in the intelligence" (p. 117).

By looking further into his work, it can found in *Darwin and The Humanities* in which Baldwin (1909) presented that "in cases where the intelligent or other adjustive factor is on the whole of greater utility, variations towards the disintegration of the instinctive congenital part, would be selected. The growth of intelligent action superseding instinctive" (p. 21), and that "once admit that the intelligence, even in its simplest forms, as seen in imitation, play and the resulting accommodative actions, can be applied to the learning of anything, and that variations in plasticity are selected to allow of its development this once admitted, we have the possibility of a continuous handing down from generation to generation, a Social Heredity, which is no longer subject to the limitations set upon physical heredity" (p. 28).

Here it seems to us that for Baldwin, with social heredity, there is **no need of fixing phylogenetic variation for previously acquired behaviour** if organisms can easily acquire those behaviour through imitation, teaching, or just copying.

B. The Baldwin-Simpson Effect:

There have been quite a few reasons why the Baldwin's idea was not common in the literature of both psychology and biology. I do not want to go too far here, yet one of those was *the Baltimore scandal* in which the head of psychology department of Johns Hopkins University (Baldwin) was caught by the police, which then made him mostly disappear from any scientific community (Horley (2001)).

Baldwin's original idea of organic selection seemed to come back to scientific discourse through Simpson (1953), appeared in Evolution 1953, in which the idea was first called the *Baldwin Effect*. Interestingly, Simpson's interpretation of the Baldwin Effect seemed to be stimulated by the idea of *genetic assimilation* by Waddington (1953) in the same issue. We shall call this the Baldwin-Simpson Effect since it has some differences from the original version.

Through Simpson's interpretation, the Baldwin Effect (or the BS Effect) occurs in two phases: Phase 1, individuals that through lifetime learning acquire an adaptive behaviour needed for the survival in its current environment occupy the population; and Phase 2, then the evolutionary path finds the innate trait that can replace the learned trait because of the *cost of individual learning*. Phase 2 was conflated with the idea of genetic assimilation of acquired characters by Waddington (1953) in his experiments to study epigenetics with drosophila. Interestingly and ironically, Simpson (1953) gave birth to the catchy name of the effect just for the intention of deflating the interest in the Baldwin Effect. Simpson was skeptical of the Baldwin effect as he posited that if learned behaviors do become genetically underwritten, a population will favour long-term fixed adaptation at the cost of short term and more plastic [learned behaviors], thus corrupting the point of the Baldwin effect. By the early sixties, a deeper skepticism came from a famous figure in evolutionary theory, Mayr (1963), which was then followed by Dobzhansky (1970). These authors all disagreed with Phase 2 of the Baldwin Effect (or BS effect) as evolution should favour plastic phenotype, rather than collapsing *norms of reaction* for fixity.

C. The Baldwin-Simpson Effect in Computation

The Baldwin Effect gradually gained more attention since the classic and elegant computational model by Hinton and Nowlan (1987) (henceforth H&N). H&N used the same metaphor as Simpson and attempted to demonstrate that the Baldwin Effect (or the BS effect) can occur. Figure 1 describes the detail of a replication of H&N's model which results in the same conclusion. The result from Hinton and Nowlan (1987) did stimulate the doyen of British biologists

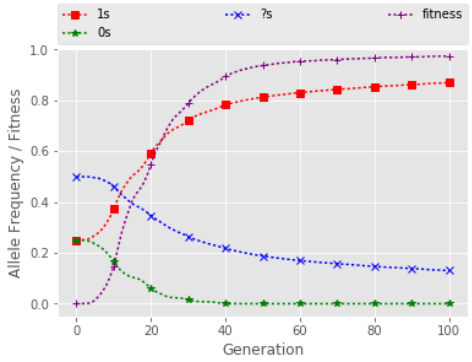

Figure 1: Replication of H&N's experiment. The task is to find the all-ones target string 111…1 (20 bits). There is only one correct solution, the target string, which has the fitness of 20. All other configurations are wrong and have the same fitness of 1. This forms a Needle-in-a-haystack landscape whereby an evolutionary search alone cannot find the solution. H&N used a different encoding. A genotype now is intialised with 3 alleles: 25% 0, 25% 1, and 50% ?. The plastic allele ? allows for lifetime learning(or plasticity), over 300 rounds (since the H&N's original 1000 was often criticised as too big by many). On each round, an individual agent is allowed to perform individual learning by changing its allele ? to either 0 or 1 as the expressed value. After learning, the fitness of an individual is calculated as: 1 + 19(300-n)/300 (n is the learning trials performed to find the solution). The population consists of 1000 individuals, crossover is only the genetic operator employed, and selection is based on fitness-proportionate as in Hinton and Nowlan (1987). We run the simulation through 100 generations, and over 30 independent runs. The frequency of the allele is plotted against the average fitness normalised in [0, 1]. There is small difference in detail perhaps due to different programming environments, yet the overall trend is the same with the original model. The Baldwin-like Effect is claimed as the frequency of 0 disappears, the frequency of correct allele 1 is increased (also the average fitness), and the frequency of plastic allele decreases as an instance of genetic assimilation due to the cost of individual learning.

Maynard Smith (1987) to feature "when learning guides evolution" in Nature Magazine. Dennett (1991) adopted the same idea to explain consciousness.

The model developed by Hinton and Nowlan, though simple, is interesting, as it opens up the trend followed by a number of studies investigating the Baldwin Effect, or how learning affects evolution, in the computer, including Mayley (1996), Harvey (1996), Mayley (1997), Suzuki and Arita (2007)). These studies interpret the Baldwin Effect in two phases, and stress the importance of the assimilation phase. Mayley (1997) and Mayley (1996) studied quite thoroughly how the cost-benefit trade-off of individual learning that could trigger genetic assimilation. Interestingly, the H&N's model has been criticised that it could not reach the state when the whole adaptive behaviour (all-ones) is assimilated, leaving no plasticity (Harvey (1996), Santos et al. (2015)). The so-called *effect* has also been employed in artificial intelligence, yet the goal is to to borrow phenomena of evolution and learning (even social learning) to create more intelligent agents to solve a problem of interest, rather than understanding the Baldwin Effect (Le (2019), Le et al. (2019)). All of these studies, for or against the effect, rely on the reinterpretation of Simpson, or the BS Effect.

D. The Recovery in Modern-day Interest

More than a century later, the ideas set out by Baldwin have also been recovered in other fields such as Evo-Devo (West-

Eberhard (2003)), Cognitive Science (Dennett (1991)). Especially, in an edited book by Weber and Depew (2003), present-day discussions about the Baldwin effect from different points of view, including epigenetics, language evolution, niche construction theory (Odling-Smee et al. (2003)) are presented. The Baldwin's 1986 paper was also cited in the recent movement in Evolutionary Biology, called the *Extended Evolutionary Synthesis* (EES) (Laland et al. (2015), Pigliucci (2007)), which tries to incorporate many factors, including epigenetics and developmental processes (West-Eberhard (2003)), in evolution that have been neglected for years in the mainstream evolutionary biology. I shall not be going too far at this moment yet it can be seen that the Baldwin Effect, which emphasises the active role of intelligence or phenotypic plasticity in evolution, can fit into, and even somehow boost the active status of the EES framework.

However, many of them are still not so clear whether the Baldwin effect requires the need for acquired characters to be assimilated. West-Eberhard (2003) says that "Baldwin conceived of it (organic selection) as a mechanism that could, in principle, lead to the reduction of plasticity as the trait in question comes under increasingly powerful genetic influence. Yet this stands at odds with the remarkable flexibility exhibited by observed organisms". The whole book dedicated for the reconsideration of the Baldwin effect by Weber and Depew (2003) also presents the controversy within the selected authors in that edition on the issue of genetic assimilation, which has led to an even stronger skepticism of what the Baldwin Effect really is, as reviewed by Sterelny (2004) and Shettleworth (2004). Shettleworth even concluded her review by referring to Depew, saying that there is really no such thing as the Baldwin Effect.

Paradoxically, what is missing from the majority of the available bibliography is the original viewpoint from which Baldwin actually formed his theory of organic selection and social heredity. Most of the contemporary discussions on the Baldwin effect seem to rely on the Simpson's interpretation. As we have argued so far, Baldwin's original factor in evolution can argue that organic selection can drive greater plasticity, escaping from genetic assimilation.

E. Concluding Remarks

Now we can feel at ease to conclude that originally Baldwin stressed on the importance of intelligence, which includes ontogenetic learning as a form of phenotypic plasticity, in directing evolution. He was right to say that the future evolution will follow the path laid by what adaptive behaviour has been acquired before. Indeed, social heredity should not be neglected when studying the "effect" on evolution.

We can offer another important point here. It was the reinterpretation of Simpson that conflated the Baldwin Effect with the idea of genetic assimilation that has raised a strong skepticism of the effect. This interpretation has had a relatively strong influence on the study of the Baldwin Effect in many disciplines, including ALife. This, indeed, restricts the original idea of the Baldwin-Baldwin Effect. Moreover, it is the lack of social heredity in the Baldwin-Simpson Effect that made the skepticism even stronger. What has been shown informs us that there exists a scenario, with the presence of social heredity, in which the Baldwin Effect occurs differently from the genetic assimilation process as often believed previously, promoting more plasticity to facilitate future intelligent acquisitions by learning.

In the next section I briefly present what Baldwin thought of social heredity and its relationship to the contemporary research on social learning and cultural evolution. I then describe the experiment to study the Baldwin-Baldwin Effect through the prism of social heredity.

II. The Baldwin Effect through Social Heredity

A. Social Heredity

Baldwin proposed social heredity as an important inheritance mechanism in which cultural knowledge and values can be transmitted both within and between generations. Baldwin (1909) said that "when we come to ask for a full account of the propagation of mental acquisitions from generation to generation, we find it necessary to recognise another form of handing down or real transmission" (p. 28). In Mental Development, Baldwin described social heredity as largely independent of physical heredity. However, Baldwin (1896), Baldwin (1902), and Baldwin (1909) later acknowledged that the two modes of inheritance can interact and have influence on each other. Baldwin (1902) wrote that "social heredity keeps certain variations alive, thus sets the direction of ontogenetic accommodation thereby influences the direction of the available congenital variations of the next generation, and so determines phylogenetic evolution" (p. 103).

Interestingly, what Baldwin once proposed more than 100 years ago bears a flavour similar to the so-called *gene-culture coevolution*, or *dual-inheritance theory*, currently promoted by cultural evolution researchers, such as Peter J. Richerson (2005), Lumsden and Wilson (2005). Gene and culture are said to co-evolve to further adaptivity of social or cultural species. Learning, both asocial (individual) and social, are media to trigger the establishment and transmission of cultural adaptations. More interestingly, the cost-benefit relationship between social learning (SL) and individual learning (IL) can produce variable evolutionary dynamics (Laland (2018), Peter J. Richerson (2005)). A combination of both trial-and-error and imitation learning is often said to produce more adaptivity, especially in human cultural evolution (Peter J. Richerson (2005)). Importantly, culture has been said to emerge only when the fidelity of cultural transmission is high (Laland (2018)). We shall incorporate fidelity of cultural transmission in our experiments in the next section.

B. Experiments and Results

In this section I present a simple computer simulation as an extension of H&N's replication, combining evolution, individual learning, and cultural inheritance. Cultural inheritance here is understood as the transmission of behaviour from parents to their offspring, vertically via social, or imitation, learning. Some limitation on this computational model should be noted. First, as previously shown in Harvey (1996), Mayley (1996), the H&N's landscape is extreme. Individual learning is quite random. Importantly, it was mostly criticised as it cannot lead to the absolute assimilation of the correct behaviour (all-ones), thus it is not the Baldwin Effect (Santos et al. (2015)). However, as I have shown in the theory part, the Baldwin's original effect does not necessarily mean the assimilation of acquired characters is required. Indeed, we shall being seeing the reverse.

For that reason, we can feel at ease to replicate the elegant H&N's model. For now, it is the transparent simplicity of H&Ns original work which is critical to its impact; such simplicity is our preference while adding new mechanisms to study the effect of interest by two experimental setups.

B1. Setup I: Evolution with Social Learning alone I propose the social learning procedure via imitation as described in Algorithm 1 below. The imitative process works as follows: For each question mark position, the observer will decide whether to copy exactly the $trait$ or a mutated version of that $trait$ from the demonstrator based on the parameter $fidelity$ which represents the fidelity of the social transmission. By default, the $fidelity$ is set to 1, that means imitative process will copy exactly the values from the demonstrator to the observer.

Algorithm 1 IMITATION

```
function IMITATION(observer, demon, fidelity = 1)
    questions = [] comment: question mark array
    for position i ∈ observer.pheno do
        if i =? then
            questions.add(i)
            observer.learning_attempt += 1
        end if
    end for
    for i ∈ questions do
        if rand() < fidelity then
            observer.pheno(i) = demon.pheno(i)
        else
            observer.pheno(i) = 1 − demon.pheno(i)
        end if
    end for
end function
```

Algorithm 2 presents the process in which evolution is combined with only social learning in place of asocial learning as in H&N's model (denoted by EVO+SL). The demonstrator is set to be the better parent of an individual. This represents a vertical cultural inheritance process, as described above. After social learning, the population operates an evolutionary process as in H&N's model described in Figure 1.

Algorithm 2 EVO + SL

```
function EVO+SL(pop, fidelity = 1) comment: Do life-time learning
    for ind ∈ pop do
        demon = ind.max_parent() comment: extract the better parent
        Imitation(ind, demon, fidelity) comment: do imitation
    end for
    comment: Evolve the population
    Do selection, reproduction, replacement
end function
```

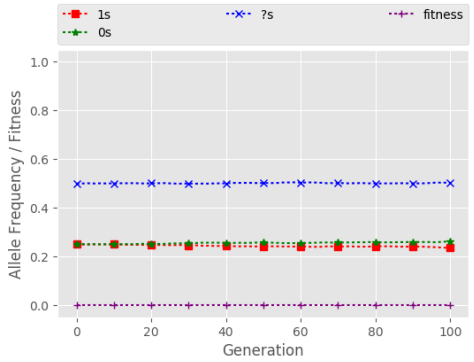

Figure 2: EVO+SL alone. Fitness is normalised in [0,1]

Look at the result in figure 2, without individual learning, social learning fails to guide evolution in the H&N's landscape. The Baldwin effect does not show up in this case. Figure 2 shows that frequency of all three alleles keeps relatively constant. No individual can find the solution, as shown in the lowest average fitness.

It is not hard to explain this. SL is *information-parasitism* – can only learn from information, or solution, produced by others. The H&N's landscape is quite special in this case. Without individual learning, there is no gradient for evolution to seek for the solution. In other words, without the presence of individual learning, no solution will be found in the evolving population. All evolving individuals are wrong. Social learners that copy from their wrong parents become wrong. Simply speaking, social learning cannot learn anything that has not been learned.

There is no influence of organic selection on evolution in this case, hence no Baldwin-Baldwin Effect.

B2. Setup 2.2: Evolution + IL + SL

Based on the analysis above, we design Algorithm 3 combining evolution with both social and asocial learning, or evolution with a *learning strategy*. A strategy is set as at each generation, an agent performs social learning based on Algorithm 1 only when its demonstrator is correct, otherwise the agent seeks for the solution individually. The demonstrator of an agent is again the better individual amongst its parents. The demonstrator is said to be correct when its fitness value is greater than 1. This is because 1 is the lowest fitness in our landscape, and an agent has its

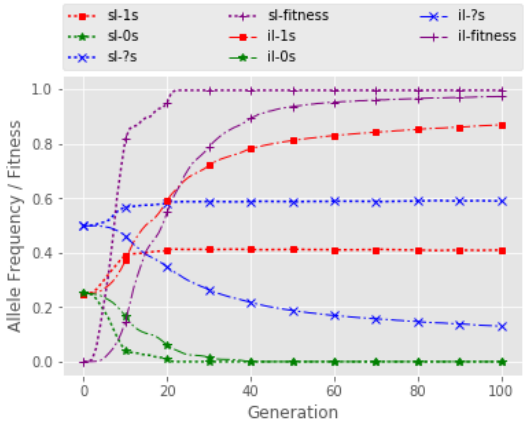

Figure 3: EVO+IL+SL vs EVO+IL. sl = EVO+IL+SL, il=EVO+IL.

fitness greater than 1 only when it successfully found the solution. After this lifetime learning process, the population goes through selection and reproduction.

Algorithm 3 EVO+IL+SL

function EVO+SL(pop, $fidelity = 1$) **comment:** Do life-time learning
  for $ind \in pop$ **do**
    $demon = ind.max\_parent()$ **comment:** extract the better parent
    if $demon.fitness > 1$ **then**
      $Imitation(ind, demon, fidelity)$
  else
      $ind.individual\_learning()$
    end if
  end for
  comment: Evolve the population
  Do selection, reproduction, replacement
end function

In Figure 3, we plot our EVO+IL+SL against the H&N's setup (EVO+IL) to see the difference between the two "*effects*". It is shown that social learning in combination with asocial learning can also *direct* the underlying evolutionary process. More specifically, we see that the frequency of wrong allele (0s) drops to zero quicker in EVO+IL+SL (at around generation 20). Contrary to the effect found in EVO+IL, EVO+IL+SL maintains a higher proportion of plasticity than the correct allele (1s). After generation 20, all the alleles in EVO+IL+SL relatively keep constant. This means there is no pressure to replace the plasticity with the fixation of $1s$. Also, the average fitness of EVO+IL+SL reaches the higher point and sooner than that of EVO+IL.

How the Baldwin Effect can be interpreted here? We observe that the behaviour of EVO+IL+SL can be divided into two phases: In the first phase, which includes 20 first generations, through individual learning some agent can find the solution. That successful agent should have no $0s$ in its genetic composition at first, and will be favoured by selection, leaving more offspring, promoting its allele configuration (with $1s$ and $?s$) in later generations. Moreover, the offspring of successful agents (without $0s$) tends to have its genotype consisting of no $0s$. Since its parent now is successful, via **social heredity** that offspring can **copy** the successful behaviour from its parent, and becomes successful too. Its genetic makeup will also be promoted, without $0s$. Thus the proportion of 0 will quickly diminish.

In the second phase, we observe that there is relatively no change in frequency of 1s and ?s, and the average fitness reaches its highest point. The explanation for the observation here is that once the frequency of $0s$ is zero, every individual in the population will have only $1s$ and $?s$ in its genotype. Each individual agent now has a chance to be successful via individual or social learning. We call it *potential agent* from now. Moreover, once the correct solution is found (in previous generations), the cultural inheritance as a vertical transmission will transmit the correct behaviour down to generations very quickly since the potential learner can copy **exactly** the solution yet with little learning attempt (the nature of our imitation algorithm). The fitness function as depicted in Figure 1 says that a lower learning cost results in a higher fitness for the learner. Therefore, the average fitness of the population in our Evo+IL+SL is higher than that in EVO+IL. That also indicates that having more plastic alleles, specifying the ability to learn socially, is more adaptive in the future, hence the dominance of '$?s$'.

Information Fidelity

One notable factor in the explanation above is the ability to transmit **exactly** the solution down to later generations. I argue that the default $fidelity = 1$ makes it much easier for the child to copy the correct solution with the much less cost. This indicates that the information fidelity could have an influential role on the effect of social heredity on evolution. This argument should be validated by running EVO+IL+SL with different levels of fidelity. For example, here we choose 0.8 and 0.5. One interesting thing is that when $fidelity = 0.5$ the imitation process as shown in Algorithm 1 performs pretty much the same as a random guessing. This is because a plasticity '?' now, on average, has 50 percent of being correct as '1', or incorrect as '0'. Thus, it is highly expected that the behavior of social learning when $fidelity = 0.5$ is quite similar to that of individual learning alone as in H&N's simulation.

In Figure 4 and 5, it can be observed that the higher the fidelity, the higher the plastic allele, the less the amount of '1', the higher the average fitness, and vice versa.

Particularly, when $fidelity = 0.5$, there is little difference in performance between EVO+IL+SL and EVO+IL in all criteria. The results obtained here are as expected and also consistent with what we have argued so far.

An explanation for this can be through the cost-benefit of social learning. When the fidelity is high, a potential agent by imitation tends to spend less learning effort than it does by trial-and-error. This leads to the fact that an agent having more plastic alleles has a higher average fitness. The selection process will favor that kind of plastic allele over others.

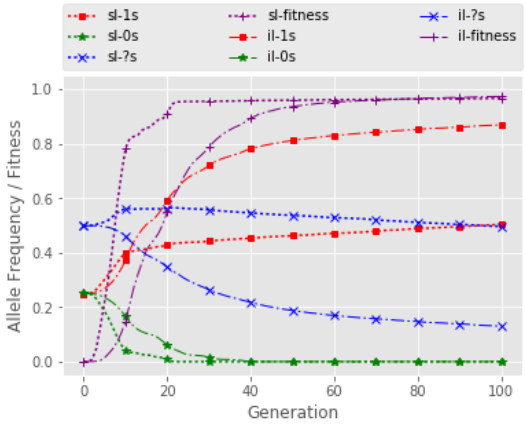

Figure 4: EVO+IL+SL vs EVO+IL. Fidelity = 0.8.

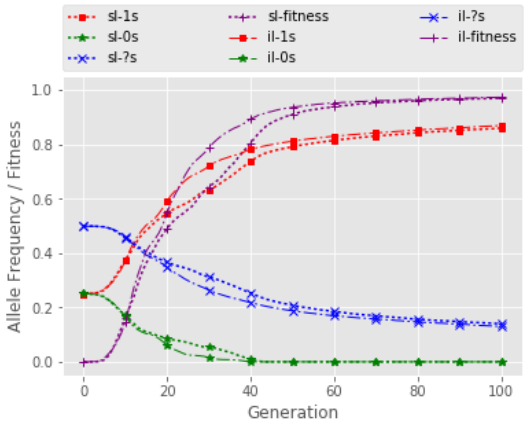

Figure 5: EVO+IL+SL vs EVO+IL. Fidelity = 0.5.

When the fidelity decreases, an observer has more chances of not copying correct values from the demonstrator. This means some plasticity '?' results in higher chance of being incorrect (having the value of 0). Now having more plasticity '?' means having more possibility of being incorrect. This also means that each plastic value in this case requires more learning effort to find the correct value of 1. Thus, having fewer number of '?' reduces the learning cost. Again, the selection process will favor a correct individual with less learning cost, the allele '?' will be less favored when the fidelity is lower.

From all of the observation and analyses above, we can conclude that information fidelity plays an important role in how social heredity directs evolution.

Conclusion and Further Discussion

In this paper, I have reconsidered the Baldwin Effect in both theoretical and empirical (computational) points of view. By briefly discussing the literature of interest, I have shown that Baldwin did not restrict the effect to genetic assimilation – which has mostly been used to understand the Baldwin Effect for many years in trans-disciplinary discourse, including in computational studies. What is implied here is that the Baldwin Effect should not be conflated with the idea of genetic assimilation, instead genetic assimilation may just be one of the ways through which the Baldwin Effect may occur. Social heredity has also been shown to play an important role in directing evolution.

Experimental results support what has been theorised. Through a specific landscape and parameter settings, it has been empirically shown that without individual learning, social heredity shows no "effect" at all. This shows that the adaptive behaviour should exist first, before social heredity takes place. When coupled with individual learning, social heredity via social learning can direct evolution in different ways depending on the fidelity of the cultural transmission. When the fidelity is high, plasticity is promoted more than the assimilation of acquired characters; yet when fidelity goes down, more assimilation emerges.

Here and now I would like to pose a question that why we should be, and keep being, interested in the Baldwin Effect. It seems that this question should have been mentioned earlier. Yet I think that only after we have presented and explained the effect in Baldwin's original spirit and how it differs from what has often been understood, it is less uncertainty to talk about what the original Baldwin Effect, or the Baldwin Effect, would imply. One plausible reason, to me, is that the effect, if happens, helps explain why and how evolution can be directed by intelligent faculties which are also the products of evolution. This stresses the role and importance of intelligence, mind, behaviour, or any form of ontogenetic development in evolution. This also means there are circumstances in which the phenotype is not just the passive product of the gene and environment, but plays an active role in directing the evolutionary pathway of the species. The Baldwin Effect, I think, implies a **reciprocal causation** in evolution that phylogeny and ontogeny should be considered both causes and consequences. This line of thought can change the way we understand and explain evolution in biological, cultural, and even *artificial* worlds.

In the modern-day discussion of evolution there has been a call for an extension and expansion of Darwinian account of evolution via the modern synthesis (Pigliucci (2007)). Proponents of the extended evolutionary synthesis also stress the constructive role of the organism, or what we call niche construction (Odling-Smee et al. (2003)), and its reciprocal causation in evolution. This research programme has raised serious questions about the *reductionist* approach dominant in the modern synthesis, saying that not everything can be reduced to the gene (Laland et al. (2015)). Interestingly, what I have argued so far tells us that Baldwin's legacy seemed to prepare a new movement for Darwinian evolution more than 100 years ago, yet was largely neglected in evolutionary discourse for a couple of reasons in the 20th century.

For another reason, I believe the Baldwin Effect could

contribute to the explanatory repertoire of the evolution of intelligent faculties in animals, including the human mind. Baldwin (1909) once tried to link the explanatory repertoire between disciplines, from evolutionary biology to psychology to the humanities, through his ideas of organic selection and social heredity. If the Baldwin Effect occurs through human cultural niche construction processes, this can help explain how the human brain evolved to be better at learning in the changing cultural world. The role of organic selection and social heredity in evolution is also believed to have a further value in explaining the evolution of gregarious habits and cooperative behaviour in social animals. Future work will delve deeper into these avenues of research.

References

Baldwin, J. (1895). *Mental Development in the Child and the Race: Methods and Processes*. Macmillan.

Baldwin, J. (1909). *Darwin and the Humanities*. Library of genetic science and philosophy. Review Publishing Company.

Baldwin, J. M. (1896). A new factor in evolution. *The American Naturalist*, 30(354):441–451.

Baldwin, J. M. (1902). *Development and evolution, including psychophysical evolution, evolution by orthoplasy, and the theory of genetic modes.* MacMillan Co.

Dennett, D. C. (1991). *Consciousness Explained*. Back Bay Books.

Dobzhansky, T. (1970). *Genetics of the Evolutionary Process*. Columbia University Press.

Harvey, I. (1996). Is there another new factor in evolution? *Evol. Comput.*, 4(3):313–329.

Hinton, G. E. and Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems*, 1:495–502.

Horley, J. (2001). After the baltimore affair: James mark baldwin's life and work, 1908-1934. *History of Psychology*, 4(1):24–33.

Huxley, J. (1942). *Evolution, the Modern Synthesis*. G. Allen & Unwin Limited.

Laland, K. (2018). *Darwin's Unfinished Symphony: How Culture Made the Human Mind*. Princeton University Press.

Laland, K. N., Uller, T., Feldman, M. W., Sterelny, K., Müller, G. B., Moczek, A., Jablonka, E., and Odling-Smee, J. (2015). The extended evolutionary synthesis: its structure, assumptions and predictions. *Proceedings of the Royal Society B: Biological Sciences*, 282(1813):20151019.

Le, N. (2019). Evolving self-taught neural networks: The baldwin effect and the emergence of intelligence. In *2019 AISB Annual Convention – 10th Symposium on AI & Games*, Falmouth, UK.

Le, N., Brabazon, A., and O'Neill, M. (2019). The evolution of self-taught neural networks in a multi-agent environment. In Kaufmann, P. and Castillo, P. A., editors, *Applications of Evolutionary Computation*, pages 457–472, Cham. Springer International Publishing.

Lumsden, C. J. and Wilson, E. O. (2005). *Genes, Mind, and Culture*. WORLD SCIENTIFIC.

Mayley, G. (1996). Landscapes, learning costs, and genetic assimilation. *Evolutionary Computation*, 4(3):213–234.

Mayley, G. (1997). Guiding or hiding: Explorations into the effects of learning on the rate of evolution. In *In Proceedings of the Fourth European Conference on Artificial Life*, pages 135–144. MIT Press.

Mayr, E. (1963). *Animal species and evolution*. Belknap Press. Belknap Press of Harvard University Press.

Morgan, C. L. (1896). On modification and variation. *Science*, 4(99):733–740.

Odling-Smee, F., Laland, K., Feldman, M., and Feldman, M. (2003). *Niche Construction: The Neglected Process in Evolution*. Monographs in Population Biology. Princeton University Press.

Osborn, H. F. (1896). Ontogenetic and phylogenetic variation. *Science*, 4(100):786–789.

Peter J. Richerson, R. B. (2005). *Not By Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press, 1 edition.

Pigliucci, M. (2007). Do we need an extended evolutionary synthesis? *Evolution*, 61(12):2743–2749.

Richards, R. J. (1989). *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior (Science and Its Conceptual Foundations)*. Science and Its Conceptual Foundations",. University of Chicago Press.

Santos, M., Szathmry, E., and Fontanari, J. F. (2015). Phenotypic plasticity, the baldwin effect, and the speeding up of evolution: The computational roots of an illusion. *Journal of Theoretical Biology*, 371:127 – 136.

Shettleworth, S. J. (2004). Book review: Evolution and learning: The baldwin effect reconsidered. *Evolutionary Psychology*, 2(1):147470490400200.

Simpson, G. G. (1953). The baldwin effect. *Evolution*, 7(2):110.

Smith, J. M. (1987). When learning guides evolution. *Nature*, 329(6142):761–762.

Sterelny, K. (2004). A review of evolution and learning: the baldwin effect reconsidered edited by bruce weber and david depew. *Evolution and Development*, 6(4):295–300.

Suzuki, R. and Arita, T. (2007). Repeated occurrences of the baldwin effect can guide evolution on rugged fitness landscapes. In *2007 IEEE Symposium on Artificial Life*. IEEE.

Waddington, C. H. (1953). Genetic assimilation of an acquired character. *Evolution*, 7(2):118–126.

Weber, B. and Depew, D. (2003). *Evolution and Learning: The Baldwin Effect Reconsidered*. A Bradford book. MIT Press.

Weismann, A., Parker, W., and Rönnfeldt, H. (1893). *The Germplasm: A Theory of Heredity*. Contemporary science series. Scribner's.

West-Eberhard, M. (2003). *Developmental Plasticity and Evolution*. Developmental Plasticity and Evolution. OUP USA.