Cognition 197 (2020) 104165

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

What the Baldwin Effect affects depends on the nature of plasticity

Thomas J.H. Morgan[a,b,c,*], Jordan W. Suchow[c], Thomas L. Griffiths[c]

[a] Arizona State University, School of Human Evolution and Social Change, 900 S. Cady Mall, Tempe 85287, United States of America
[b] Arizona State University, Institute of Human Origins, 951 S. Cady Mall, Tempe 85287, United States of America
[c] University of California, Berkeley, Department of Psychology, Tolman Hall, Berkeley 94720, United States of America

ARTICLE INFO

Keywords:
Baldwin effect
Learning
Plasticity
Evolution

ABSTRACT

In a process known as the Baldwin Effect, developmental plasticity, such as learning, has been argued to accelerate the biological evolution of high-fitness traits, including language and complex intelligence. Here we investigate the evolutionary consequences of developmental plasticity by asking which aspects of a plastic trait are the focus of genetic change. The aspects we consider are: (*i*) dependencies between elements of a trait, (*ii*) the importance of each element to fitness, and (*iii*) the difficulty of acquiring each element through plasticity. We also explore (*iv*) how cultural inheritance changes the relationship between plasticity and genetic change. We find that evolution by natural selection preferentially fixes elements that are depended upon by others, important to fitness, or difficult to acquire through plasticity, but that cultural inheritance can suppress and even reverse genetic change. We replicate some of these effects in experimental evolutionary simulations with human learners. We conclude that what the Baldwin Effect affects depends upon the mechanism of plasticity, which for behavior and cognition includes the psychology of learning.

1. Introduction

The interaction between developmental plasticity, expressed through processes such as learning, and biological evolution has received considerable attention over the past few decades (Gabora, 2008; Scheiner, 1993; Via et al., 1995). However there remains considerable debate over whether plasticity is a driving force in evolution, taking the lead in adaptation with genetic change following afterwards (Laland, Wray, & Hoekstra, 2014; Pigliucci, Murren, & Schlichting, 2006; West-Eberhard, 2003). One way in which plasticity and evolution may interact is the Baldwin Effect (Baldwin, 1896; Weber & Depew, 2003), which proposes that evolution proceeds by selection favoring genetic variants that support adaptive traits that arise via plasticity. Over time, accumulated genetic change allows traits to be reliably acquired by all members of the species and may reduce their plasticity. This can be contrasted with an aplastic model in which novel traits arise purely through genetic mutation.

Baldwin's theory, originally called “organic selection”, has had variable success since its proposal by several figures in the late 19th century (Baldwin, 1896; Lloyd Morgan, 1896; Morgan & Harris, 2015; Osborne, 1896; Richards, 1987; Weber & Depew, 2003). Initially regarded as an important part of the evolutionary process, a scandal concerning Baldwin's personal life, the development of mathematical evolutionary theory, and an increasing segregation between developmental and evolutionary biology all contributed to Baldwin's theory falling out of favor in the early 20th century (Richards, 1987). Simpson briefly mentioned it shortly after the formation of the Modern Synthesis, coining the term “Baldwin Effect” (Simpson, 1953), but it remained on the periphery of evolutionary thinking.

Despite this mixed history, and although the Baldwin Effect remains contested (Richards, 1987; Weber & Depew, 2003), it has made a modest resurgence in the last few decades. In particular, it has been invoked in several cases that are a challenge for aplastic models of evolution, such as when fitness landscapes are heavily skewed with only a small number of genotypes associated with high fitness (Hinton & Nowlan, 1987). This *needle-in-a-haystack* problem is hard for selection to solve without plasticity because the multitude of low fitness genotypes create a flat fitness landscape. In a demonstration that has come to be associated with the Baldwin Effect, Hinton and Nowlan (1987) conducted a series of simulations in which they introduced plasticity by including an allele that did not specify the phenotype, but instead prompted the organism to developmentally explore different phenotypic options. The existence of this allele not only accelerated the rate at which beneficial phenotypes were discovered, but also accelerated the evolution of the beneficial genotype because the plastic alleles created a fitness gradient that natural selection could follow.

* Corresponding author at: Arizona State University, School of Human Evolution and Social Change, 900 S. Cady Mall, Tempe 85287, United States of America.
E-mail address: thomas.j.h.morgan@asu.edu (T.J.H. Morgan).

https://doi.org/10.1016/j.cognition.2019.104165
Received 11 January 2019; Received in revised form 17 December 2019; Accepted 18 December 2019

Other work has built on these results, showing the strengths and limitations of the Baldwin Effect. For instance, while the ability of plasticity to accelerate the *appearance* of beneficial phenotypes has received support (Ancel, 1999; Fontanari & Santos, 2017; Santos, Szathmáry, & Fontanari, 2015), theory suggests it ultimately slows their genetic fixation because once traits can be reliably acquired through plasticity there is little selective advantage in further genetic change (Ancel, 1999, 2000; Fontanari & Santos, 2017). Moreover, once a beneficial phenotype has appeared, assuming it has a sufficient fitness benefit, it can spread as rapidly in a non-plastic sexually reproducing population as in a plastic population (Santos et al., 2015). Other potential limitations include the observation that plastic traits may need to be widespread before selection can overcome genetic drift and increase aplastic genetic variants that support them (Chater, Reali, & Christiansen, 2009), and that the hypothesized reduction in plasticity following plastic adaptation may be sufficiently slow that plasticity is minimally reduced before the environment is likely to change again (Scheiner, Barfield, & Holt, 2017). Nonetheless, many models of evolution support the general plausibility of a process where (*i*) adaptation is initially plastic, (*ii*) plasticity increases in response to environmental novelties, and (*iii*) environmental stability leads to the replacement of plasticity with fixed genetic influence (Ancel, 1999; Lande, 2009; Scheiner et al., 2017).

Many animal behaviors are potential candidates for such an evolutionary process. For example, archer fish propel jets of water from their mouth to capture flying insects (Schuster, 2007), shaping their mouth to focus the jets on prey at different distances (Gerullis & Schuster, 2014). Plasticity is clearly relevant to this behavior as young fish need to learn the technique by watching others (Schuster, Wöhl, Griebsch, & Klostermeier, 2006). While it is possible that this behavior evolved without plasticity, this would have required the existence of a mutation that directly caused (at least a primordial form of) the behavior. Alternatively, ancestral archer fish may have learned to target insects above the water's surface and subsequent genetic change reinforced and supported this behavior. Such a genetic response can be seen in their eyes which have evolved to accurately perceive airborne insects from underwater (Temple, Hart, Marshall, & Collin, 2010), a trait that presumably followed, not preceded, the behavior.

Other traits for which plasticity has been argued to be important are those involving coordination between individuals, most prominently language (Deacon, 1997; Pinker & Bloom, 1990). Aplastic accounts must suppose that language arose through genetic change. However, this implies that it first appeared in a single individual. As the fitness benefit of language requires more than one individual to possess it, this would hinder its spread. Plasticity has been used in two ways to solve this problem. In the more limited case, language may have first arisen via a genetic mutation, but was nonetheless able to spread between individuals due to plasticity thereby bringing fitness benefits to its users (Pinker & Bloom, 1990). Assuming the mutation enhances language acquisition, there would then be a selective pressure favoring its spread. Other theories go further, suggesting that language both arose and spread via plasticity and that genetic change was entirely in response, favoring genetic variants that supported language acquisition (i.e., the Baldwin Effect) (Deacon, 1997).

These arguments apply to all cases where phenotypes need to be coordinated across individuals, including animal communication systems, such as birdsong. While it is possible that song could have evolved without plasticity (for instance, through sensory exploitation), consistent with the Baldwin Effect (Ancel, 1999, 2000), song development is often plastic (e.g. Thorpe, 1961) and species-typical song can emerge with remarkably limited stimuli (Feher, Suzuki, Okanoya, Ljubicic, & Tchernichovski, 2014).

Another means by which plasticity has been suggested to drive evolutionary change, this time in the case of human intelligence, is as an “evolutionary crane” (Dennett, 2003). This account supposes that our ancestors used their (more limited) cognition to develop behaviors that increased their fitness and that were sufficiently hard to acquire that selection favored genetic variants supporting their acquisition. Rather than being trait specific, genetic change increased our general cognitive capacities, which led to the discovery of even more complex behaviors and thereby redoubled selection on cognition. Thus, our evolution featured a coevolution of cognition and its products that resulted in huge changes to our cognitive abilities. In support of this argument, theoretical work has found that this process will occur provided a series of increasingly complex and successful possible behaviors exists, and that it can produce evolutionary dynamics similar to the hominin archaeological record (Morgan, 2016).

While this pervious work has identified cases in which plasticity may have played an important role, it remains unclear precisely how we might expect plasticity to affect genetic change. Which genes will be targeted? To what extent will behaviors become fixed? Here, we address these questions with theoretical analyses addressing how the effect of developmental plasticity on genetic change is modulated by (*i*) how traits interact with each other to produce fitness benefits, (*ii*) how important traits are to fitness, and (*iii*) how readily traits are acquired through plasticity. We then extend these analyses to include cultural inheritance, allowing the prevalence of the traits in the population to affect the probability they are acquired through plasticity. Finally, we test the predictions resulting from our theoretical analyses in an experimental simulation with a population of human learners. Our results present a clearer picture of what the Baldwin Effect affects: traits that are difficult to acquire, important to fitness, or that are required by other traits are likely to come under increasing genetic influence. These results have significant implications for understanding the potential role of the Baldwin Effect in explaining human traits such as language and higher-level cognition.

2. Theory

2.1. The simulation framework

Based on Hinton and Nowlan (1987), we consider a population of 1000 asexual, haploid organisms whose fitness is determined by the acquisition of n traits. The acquisition of the tth trait by the ith individual is affected by a corresponding genetic locus, $G_{i,t}$. Accordingly, each organism has n genetic loci. Each locus contains one of two possible alleles: *fixed* and *plastic*. A fixed allele means that the organism is guaranteed to acquire the corresponding trait, while a plastic allele means that it acquires the trait through plasticity with probability p. This means we assume that while plasticity has no direct costs (e.g. by delaying development, or requiring costly information processing systems) we do assume it has limitations in that it only probabilistically matches environmental conditions (Dewitt, Sih, & Wilson, 1998). As such selection will always act to replace the plastic allele with the fixed allele.

Simulations start with all individuals being randomly assigned genomes (the *plastic* allele being nine times more likely than the *fixed* allele at any given locus) and proceed for a number of generations. In each generation, all individuals probabilistically acquire the traits in question (though learning, conditional gene expression, etc.). The probability that the i^{th} individual acquires the t^{th} trait is given by $\phi_{i,t}$, where:

$$\phi_{i,t} = \begin{cases} 1, & \text{if } G_{i,t} = fixed \\ p_{i,t}, & \text{if } G_{i,t} = plastic \end{cases}. \quad (1)$$

Whether or not individuals successfully acquire the t^{th} trait is given by $\tau_{i,t}$, where:

$$\tau_{i,t} = \begin{cases} 0, & \text{if trait not acquired} \\ 1, & \text{if trait acquired} \end{cases}. \quad (2)$$

Individual (absolute) fitness is a baseline value ($F_{\min}$) plus a fitness

benefit from each trait successfully acquired. The fitness benefit to the i^{th} individual from the t^{th} trait is $f_{i,t}$. Accordingly, the fitness, F, of the i^{th} individual is:

$$F_i = F_{\min} + \sum_{t=1}^{n} \tau_{i,t} f_{i,t}. \quad (3)$$

Note that (as per Hinton & Nowlan, 1987) this assumes trait acquisition is binary, with no middle-grounds or near-misses. This assumption is made for simplicity and we recognize that, in reality, trait fitness typically varies continuously. Once fitness is calculated, reproduction occurs as follows: For each of the 1000 offspring to be produced, a single parent is chosen from among the current generation of organisms. An organism is chosen to be a parent with probability proportional to its fitness, and organisms can be chosen to reproduce multiple times. Organisms with high fitness are thus likely to have multiple offspring. Offspring inherit their parent's genome, but each locus mutates with probability q, producing a *plastic* allele with probability 0.9, otherwise a *fixed* allele (note that *plastic* to *plastic* and *fixed* to *fixed* mutations are permitted). We note that mutation is heavily skewed towards plasticity, an assumption we make to best show the differential effects of selection (which acts to increase the prevalence of the *fixed* allele) across different loci. To test the robustness of our results we ran additional simulations without this bias (see supplementary material section 1), as this did not change the results qualitatively here we only present the results including biased mutation. The offspring generation then replaces the parental generation. All simulations were repeated with sexual reproduction (including recombination) instead of asexual reproduction, but this did not qualitatively change the equilibria reached and so below we present only the results of the asexual simulations (for a selection of results with sexual reproduction see the supplementary material, section 2).

The results of this baseline simulation are relatively straightforward (see Fig. 1). As per the Baldwin Effect, selection increases the proportion of *fixed* alleles in the population. However, as per Muller's ratchet (Felsenstein, 1974; Muller, 1964), mutation decreases this proportion and the population reaches an equilibrium where selection and mutation cancel out. Accordingly, increasing the mutation rate decreases the frequency of the *fixed* allele in the population (see Fig. 1a, for a brief consideration of finer changes to the mutation rate see supplementary material, section 4). Similarly, increasing the number of loci also decreases the equilibrium frequency of the fixed allele (see Fig. 1b) as there are now a greater number of loci that can mutate. The effects of $p_{i,t}$ and $f_{i,t}$ are discussed in more detail in the following sections. (Note that our discussion here focuses on the equilibria reached, but for a more detailed discussion of the dynamics over time see supplementary material, section 5).

2.2. Trait inter-dependence

Phenotypes, whether morphological or behavioral, consist of hierarchical arrangements of modular subunits (Wagner, 2014; West-Eberhard, 2003). For instance, dust-bathing by birds involves the coordinated execution of pecking, raking and scratching the bathing substrate; squatting; wing tossing; head rubbing; and full-body rolling (Hogan, 1994). This can be seen in our above model; although each individual trait is binary (being successfully acquired or not) the set of traits is itself a somewhat-continuously varying super-trait that takes 11 different fitness values depending on how many of the 10 sub-traits have been acquired. In the previous section we assumed the fitness effects of all traits were additive, but this will often not be the case. Accordingly, we now extend our model to consider cases where each trait's fitness benefit interacts with whether other traits have been acquired. The extreme (cf. Hinton & Nowlan, 1987) is total dependence, where all traits need to be successfully acquired for any fitness benefit:

$$F_i = F_{min} + \sum_{t=1}^{n} f_{i,t} \prod_{t=1}^{n} \tau_{i,t}. \quad (4)$$

In this case the super-trait itself is binary, with only a single way it can be acquired (all 10 sub-traits need to be acquired) and 1023 different ways it can fail to be acquired. However, there are an infinite number of alternative “partial” dependencies. We consider two: “mutual dependence” and “chain dependence”.

Mutual dependence is defined as follows:

$$F_i = F_{min} + (\overline{\tau}_{i,1:n})^d \sum_{t=1}^{n} f_{i,t}, \quad (5)$$

where $\overline{\tau}_{i,1:n}$ is the average value of τ_i for traits 1 to n. It assumes that all traits depend on each other, however, unlike total dependence, some of

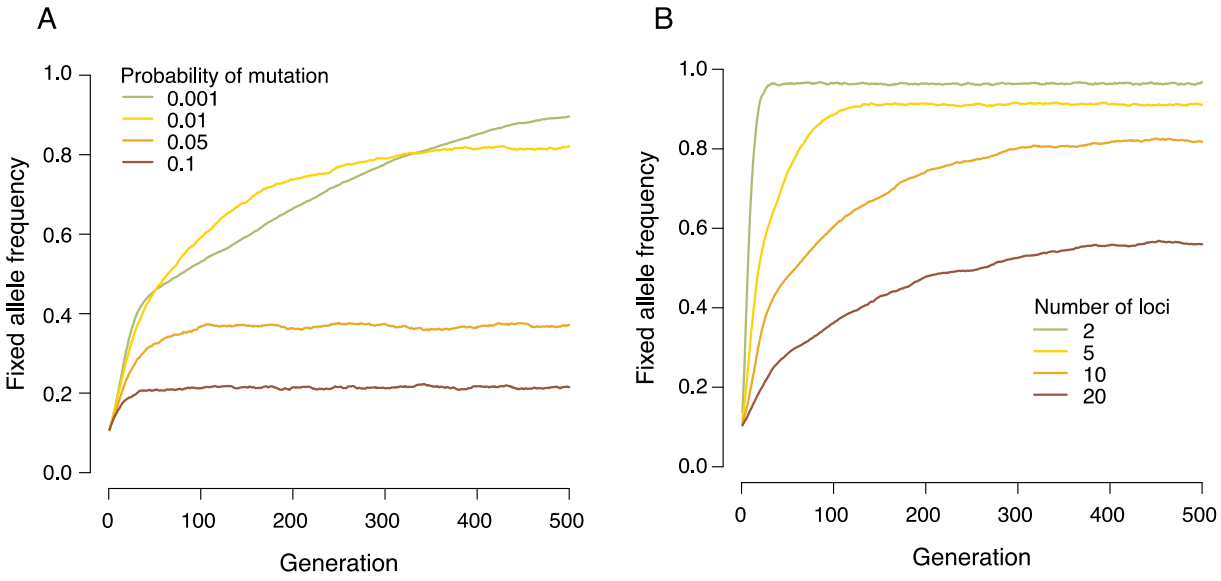

Fig. 1. The average proportion of loci with the fixed allele over the first 500 generations across 20 repeats. Parameter values are: $n = 10$, $p_{i,t} = 0.5$, $f_{i,t} = 1/n = 0.1$, $q = 0.01$ and $F_{min} = 0.01$. **A** Average fixation for different values of q – the mutation rate. When the mutation rate is higher, the equilibrium level of fixation is lower. In the cases shown, $n = 10$. **B** Average fixation for different values of n (the number of traits and loci). When there are more loci, the equilibrium level of fixation is lower. In the case shown $q = 0.01$.

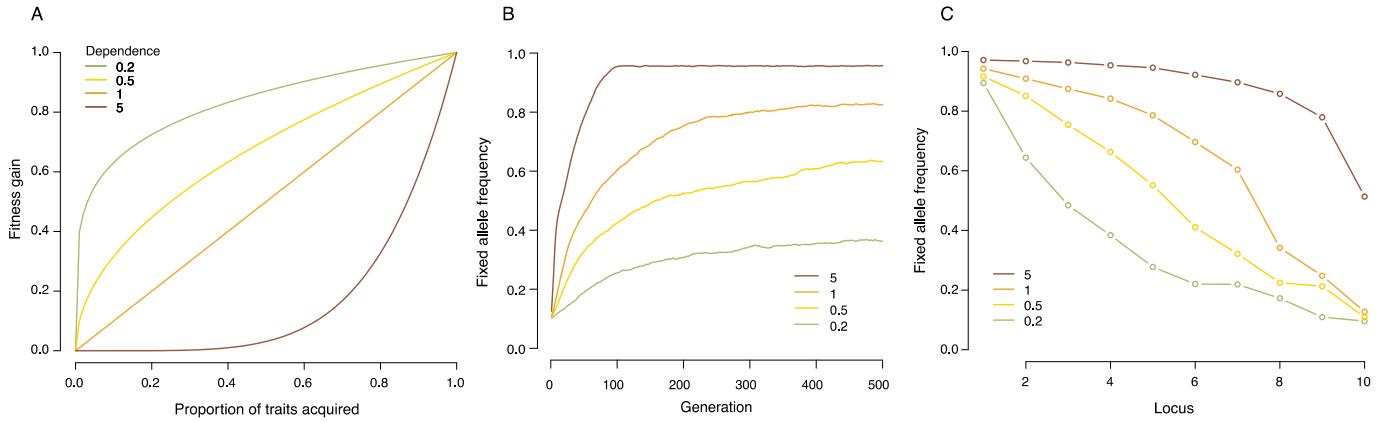

Fig. 2. The effects of inter-trait dependency on fixation. Parameter values are: $n = 10$, $q = 0.01$, $p_{i,t} = 0.5$, $f_{i,t} = 1/n$ and $F_{min} = 0.01$. **A** The relationship between the dependency parameter, d, and the payoff from acquiring a certain proportion of the available traits in the case of mutual dependency. When $d < 1$ (green and yellow lines) even a small number of traits give considerable benefit. When $d = 1$ (orange line) each trait contributes the same additive benefit. When $d > 1$ (red line) more traits are needed to receive much benefit. **B** In the case of mutual dependency, the greater the value of d the higher the equilibrium average fixation. This figure shows average fixation over the first 500 generations across 20 repeats. **C** In the case of chain dependency fixation is greater for traits that are depended on by other traits. The horizontal axis is the index of the trait and values shown are the average level of fixation for each trait after 500 generations across 20 repeats. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the potential fitness benefit can be gained even if not all the traits are acquired. How the number of traits acquired affects fitness is controlled by the dependence parameter, d, which ranges from 0 to $+\infty$ (see Fig. 2a). When d is extremely small ($0 < d < 1$), the traits exhibit interchangeability, i.e. relatively few traits are needed to gain most of the possible fitness benefit of all traits combined. In terms of the dust bathing example, this implies that dropping any one particular action (e.g. substrate raking) has minimal impact on the overall efficacy of the process provided that most other actions are present. When $d = 1$, the fitness gains from each trait are independent. As d increases beyond 1, each trait becomes increasingly important and the majority of the traits are needed to gain much benefit. As d approaches infinity, total dependence is achieved and missing even a single trait prevents any fitness gain. An example of such a critical step is the soaking of nardoo seeds by the Yandruwandha, an Australian Aboriginal group, without which the seeds are toxic (Earl & McCleary, 1994). As shown in Fig. 2b, increasing d increases average trait fixation.

The alternative implementation—chain dependence—is defined as follows:

$$F_i = F_{min} + \sum_{t=1}^{n} \left(f_{i,t}\left(\overline{t}_{i,1:t}\right)^d\right). \tag{6}$$

It assumes that the fitness benefit of the t^{th} trait is dependent only on traits 1 to (t-1). Accordingly, the first trait depends upon no other traits, but is depended upon by all other traits. Each additional trait depends upon an increasing number of other traits, but is depended upon by fewer, culminating in the final trait, which depends upon all other traits, but is depended upon by none. Again, we include a dependence parameter, d, which controls the strength of this dependency. At low values, where traits are interchangeable, having at least some of traits 1 to t will provide most of the benefit available from trait t (even if trait t itself is absent). When $d = 1$ (i.e. when traits are additive) all of traits 1 to t are needed for the maximum benefit of trait t, with each trait giving equal benefit. Finally, when d is high, missing even a single trait from 1 to t can lose most or all of the potential benefit of trait t. Examples of such processes are prepared-core techniques for the manufacture of stone tools, for instance the Levallois technique associated with the North African Middle Stone Age and Neanderthals in Europe (Adler et al., 2014; Lycett & Von Cramon-Taubadel, 2013), where raw material must first be precisely shaped before a final blow detaches the tool. As shown in Fig. 2c, the traits that are depended upon by others (i.e., early traits) show greater levels of fixation and overall fixation is increased with higher values of d.

2.3. Trait importance

We now consider how the fitness effects of a trait influence its fixation. For simplicity, we assume no dependencies between traits (i.e., each trait's fitness contribution is independent of all other traits, d is removed from the model, and fitness reverts to Eq. (3)). First, we assume that the fitness of benefit from each trait is given by

$$f_{i,t} = \frac{k}{n} \tag{7}$$

where k is a constant that controls the overall fitness value of all traits combined.

The greater the potential fitness benefit from all traits (k) the greater the overall fixation (Fig. 3a). This is caused by the traits becoming increasingly important relative to other sources of fitness (F_{min}). When k is low, the traits are a small fraction of the overall fitness of each organism and so selection acting on their fixation is weak. However, as k becomes larger, selection intensifies and the level of fixation increases. Once k is considerably larger than $F_{\min}$, further increases in k only slightly increase fixation because the effects of the traits in question already dominate the fitness of organisms.

We now allow the fitness contribution to vary across traits such that:

$$f_{i,t} = r^{t-1}\left(\frac{1-r}{1-r^n}\right) \tag{8}$$

where r ranges from 0 to 1 and controls the extent to which fitness is skewed towards a subset of the traits (see Fig. 3b). That r^{t-1} is multiplied by $(1\text{-}r)/(1\text{-}r^n)$ ensures that the overall fitness value of all traits is 1. As can be seen in Fig. 3c, traits that contribute most to fitness are fixed to a greater extent than traits that contribute less and this disparity increases with the skew among traits.

2.4. Trait difficulty

We now examine the effect of how difficult traits are to acquire through plasticity on the extent of fixation, while returning to the assumption that all traits are of equal fitness value (i.e. Eq. (7)). First, while still assuming that all traits are equally easy to acquire, we vary their probability of acquisition. Second, we assume traits vary in their difficulty of acquisition via plasticity (i.e., we allow $p_{i,t}$ to vary across t). Specifically, we consider a case where traits are evenly distributed

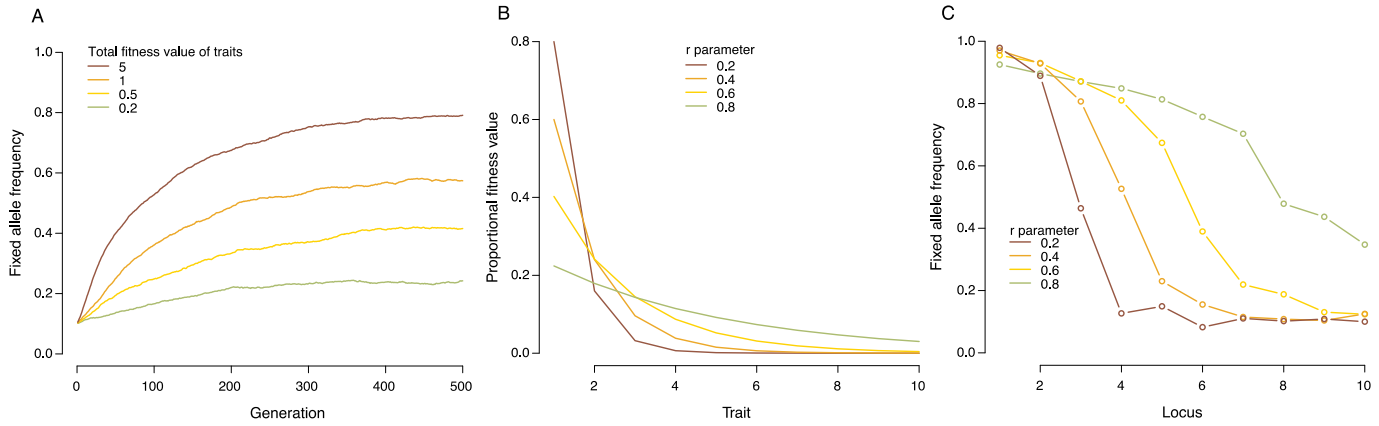

Fig. 3. The effects of trait importance on fixation. Parameter values are: $n = 10$, $q = 0.01$, $p_{i,t} = 0.5$, and $F_{min} = 1$. **A** When the traits as a whole are more important for fitness the overall fixation is higher. This figure shows average fixation over the first 500 generations across 20 repeats. **B** How the parameter r changes the fitness of each trait in the case of decreasing geometric importance. When r is high all traits contribute similarly, when r is low the first traits contribute nearly all the potential benefit. **C** When fitness is more evenly distributed among traits, fixation is also distributed more evenly across traits, however, when fitness is unevenly distributed, selection preferentially fixes the important traits. When fitness is heavily skewed most traits show little fixation, this is because they contribute so little to fitness that there is little advantage to their fixation. Nonetheless, when fitness is extremely skewed (e.g. $r = 0.2$) even loci associated with small fitness benefits (e.g. locus 2) can show high levels of fixation because they are more important than most other loci and selection can support the evolution of the fixed allele at multiple loci. The horizontal axis is the index of the trait/locus and values shown are the average level of fixation for each trait after 500 generations across 20 repeats.

between the extremes of 0 and 1.

In both cases, increasing the difficulty of acquiring traits through plasticity increases their fixation. In the case of overall difficulty, all loci are fixed to the same extent, but this level increases with difficulty (see Fig. 4a). In the case of differences between traits, the most difficult to acquire traits are fixed to greater extent than traits that are easier to acquire (see Fig. 4b). These results are consistent with other work, where giving individuals more opportunities to learn reduced the probability that beneficial aplastic alleles would reach fixation (Fontanari & Santos, 2017).

2.5. Culture

As discussed in the introduction, the Baldwin Effect has often been used in discussion of the evolution of language. However, a key difference between language and the traits that we have been modeling is that language is acquired from other individuals via social learning. As such we now include culture in our simulation by allowing the probability that a trait is acquired through plasticity to be a function of its presence in the previous generation such that:

$$p_{i,t} = p_{min} + \rho(p_{max} - p_{min}) \quad (9)$$

where $p_{\min}$ is the minimum probability of acquiring the trait, $p_{\max}$ is the corresponding maximum, and ρ is the proportion of individuals in the previous generation who acquired the trait. Note that this returns to the assumption that all traits are equally likely to be acquired through plasticity (i.e. $p_{i,t}$ is constant across t).

Fig. 5a shows the results of simulations with values of p_{min} and p_{max} ranging from 0 to 1 at intervals of 0.1. Increasing either p_{min} or p_{max} decreases the equilibrium frequency of the fixed allele, but p_{max} has a much greater influence than p_{min}—i.e., the genetic equilibrium is determined by the traits' "cultural learnability" to a greater extent than by their "asocial learnability" (further analyses, see supplementary material section 3, suggest p_{max} has roughly six times the influence of p_{min} on

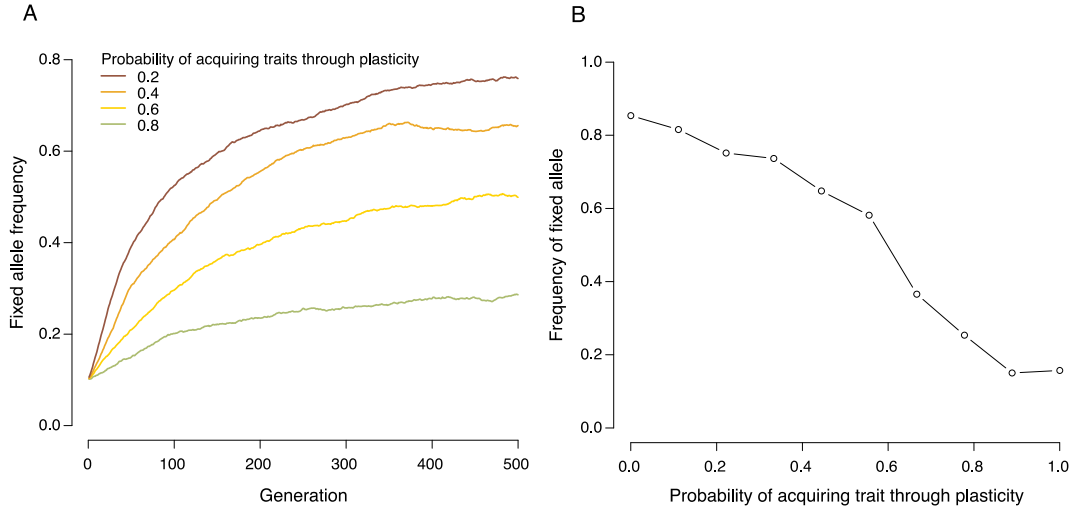

Fig. 4. The effect of trait difficulty on fixation. Parameter values are: $n = 10$, $q = 0.01$, $f_{i,t} = 1/n$ and $F_{min} = 1$. **A** The harder the traits are to acquire (i.e., the lower the probability that they are acquired if plastic) the greater the equilibrium fixation. This figure shows average fixation over the first 500 generations across 20 repeats. **B** When traits vary in the ease with which they can be acquired through plasticity, the fixation is greater for the more difficult traits than for the easier traits. The horizontal axis is the probability that each trait is acquired if plastic and the vertical axis is the average level of fixation for each trait after 500 generations across 20 repeats.

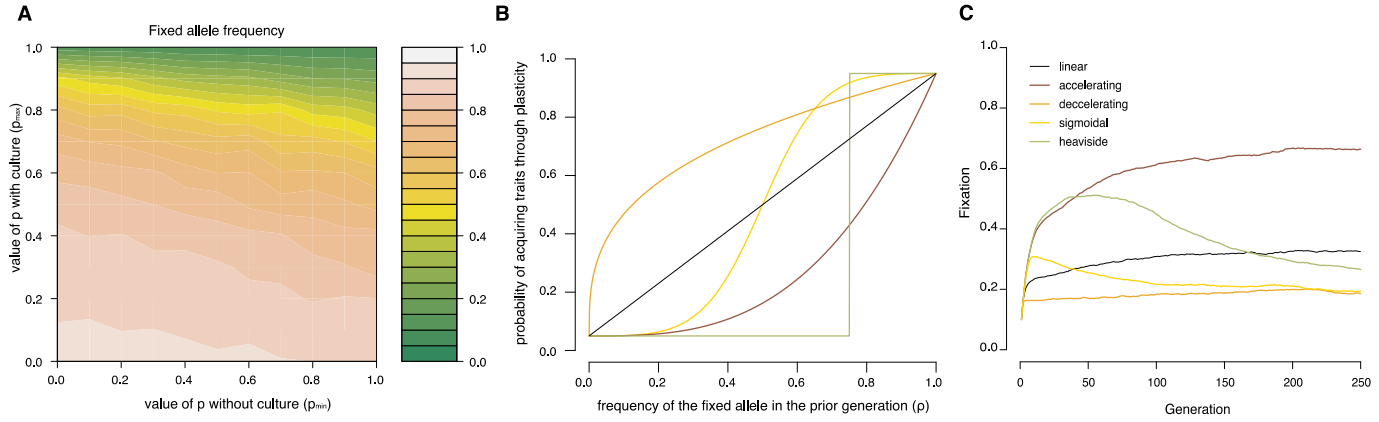

Fig. 5. The effect of culture on the Baldwin Effect. **A** When culture can affect learnability the spread of the fixed allele is largely determined by how hard a trait is to learn assuming that a large number of the population have acquired it (p_{max}, the vertical axis). This figure shows average fixation after 20 repeats of 250 generations. Parameter values are: $n = 10$, $q = 0.01$, $f_{i,t} = 1/n$ and $F_{min} = 0.01$. **B/C** The dynamics by which the equilibrium is reached depend upon the assumed relationship between culture and learnability (panel B illustrates 5 possible relationships; panel C shows the resultant dynamics assuming $p_{min} = 0.05$ and $p_{max} = 0.95$). When the relationship is linear, accelerating or decelerating the dynamics are consistent with previous models. However, if the relationship is sigmoidal or a Heaviside function (i.e. a step function) then the fixed allele initially increases, before dropping back down and reach equilibrium.

the fixed allele). To illustrate, when p_{max} is 0, the equilibrium frequency is always around 90%, while when p_{max} is 1, the equilibrium frequency is always around 10%. This is because the traits tend to increase in frequency as the simulation runs, whether by plasticity or though the evolution of fixed alleles. As such, at equilibrium, it is p_{max} that is most relevant to the efficacy of plasticity.

The evolutionary dynamics depend on the form of Eq. (9). In the case presented above, the effect of ρ on $p_{i,t}$ is linear. However, other relationships are possible (see Fig. 5b for graphical illustrations), for instance, accelerating:

$$p_{i,t} = p_{min} + \rho^3(p_{max} - p_{min}), \tag{10}$$

decelerating:

$$p_{i,t} = p_{min} + \rho^{\left(\frac{1}{3}\right)}(p_{max} - p_{min}), \tag{11}$$

sigmoidal:

$$p_{i,t} = p_{min} + \frac{\rho^3}{\rho^3 + (1-\rho)^3}(p_{max} - p_{min}), \tag{12}$$

or Heaviside (i.e. a step function):

$$p_{i,t} = \begin{cases} p_{max}, \rho \geq 0.75 \\ p_{min}, \rho < 0.75 \end{cases}. \tag{13}$$

As shown in Fig. 5c, functions where $p_{i,t}$ remains small until quite large values of ρ are reached (such as sigmoidal or Heaviside relationships) produce evolutionary dynamics where the fixed allele initially increases in frequency before returning to baseline levels.

2.6. Summary

Through a series of simulations, we have explored how the details of plasticity affect the extent of genetic fixation. We find that selection favors the genetic fixation of traits that are depended on by other traits, are difficult to acquire through plasticity, or have large fitness effects. This is true both for the set of traits as a whole, but also for individual traits; if one trait is more essential, difficult or important that others then it is preferentially fixed by selection. These results suggest that if a trait were to have evolved via the Baldwin Effect, the nature of the genetic influence over the traits development may be somewhat predictable. For instance, in the case of language, we might expect genetic evolution to support acquisition of key linguistic concepts (such as symbolism) or motor abilities (such as fine control of the tongue). However, our simulations also suggest that when cultural inheritance supports plasticity then there is less need for genetic fixation and populations can pass through a genetic transitional period followed by the reliable cultural inheritance of a plastic trait. Given that language is extensively culturally inherited, this suggests that language acquisition may be only minimally supported by genetic variants, even if this was not the case in the past.

3. Experiment

The above simulations rely on simplifying assumptions about the nature of learning and the traits in question. This means there is a need to verify whether theoretical results hold with real learning capacities which may depart from these assumptions. For instance, errors may be biased, or individuals may vary in their ability to learn or in their motivations. Therefore, to empirically test some of our theoretical findings, we conducted an "experimental evolutionary simulation" with human learners taking the place of simulated agents, and an experimental design that replicates some of the theoretical contexts considered above.

3.1. Material and methods

We carried out a single experiment involving 18 parallel experimental evolutionary simulations. In each simulation a population of 60 participants evolved for 40 generations (for a total of 2400 participants), each participant playing the role of a single agent. Each participant took part once in all 18 simulations (and in the same generation in each simulation) so while the total number of agents is 43,200, the total number of participants remains at 2400. (Note that although participants take part in multiple simulations, because each participant progresses though simulations in a random order, and remains in the same generation across simulations, practice effects are unlikely to affect results.) Participants were recruited through Amazon's Mechanical Turk, gave their consent, were given instructions, completed the experiment and were then debriefed and paid. Recruitment occurred one generation at a time: experimental slots were made available in batches of 60 (the population size) and once all 60 participants had successfully taken part in all 18 repeat simulations the next batch was recruited. This process was repeated until all 40 generations of all 18 simulations were complete. Recruitment and testing were approved by the Committee for Protection of Human Subjects at University of California, Berkeley (protocol ID 2015-12-8227). Participants were paid $1 for

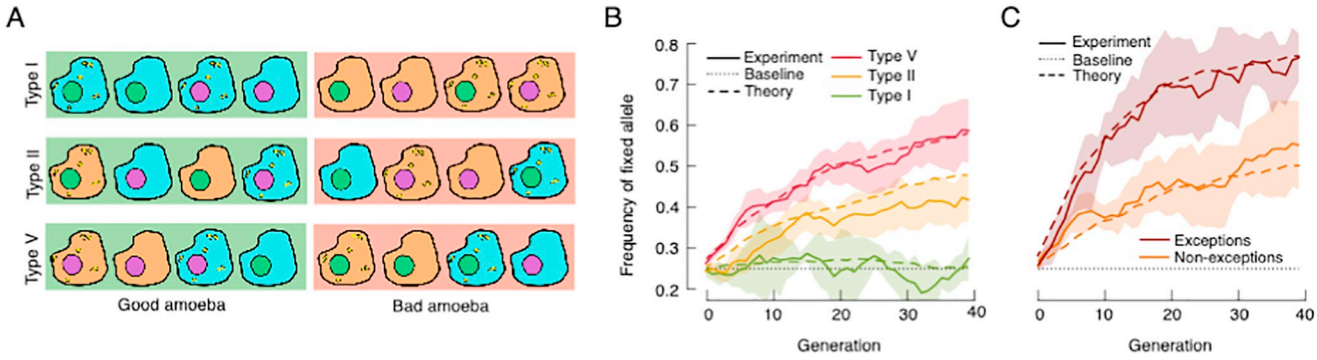

Fig. 6. Stimuli and results of the experimental simulation with human learners. **A** Three example rules for amoeba categorization. Top; a Type I rule, "blue amoebas are good". Middle; a Type II rule, "orange amoebas with green nuclei and blue amoebas with purple nuclei are good". Bottom; a Type V rule, "amoebas with purple nuclei are good unless they are blue and don't have spots in which case amoebas with green nuclei are good". **B–C** Average evolutionary dynamics across simulations (panels share a single vertical axis). Experimental results are presented as mean-of-means frequency of the *fixed* allele across repeat simulations (solid lines) ± 1.96 standard errors (shaded area, equivalent to a 95% confidence interval assuming normality). Theoretical predictions (dashed lines) and baseline expectation under mutation alone (dotted line) are included for comparison. Theoretical expectations were generated with a version of the theoretical model, modified to include the experimental fitness function, mutation rate and bias, sexual reproduction and agent performance comparable to that of the human participants. Experimental results closely match theoretical expectations: The *fixed* allele increased the most in simulations with a Type V rule. With a Type I rule it was consistent with the level expected due to mutation. Within simulations using a Type V rule the *fixed* allele was more prevalent at loci corresponding to the exception amoebas than at loci corresponding to non-exception amoebas. With a Type II rule the experimental frequency of the fixed allele was typically slightly lower than theoretical expectations. However, this deviation is modest (the theoretical expectation is within 2 standard errors of the data) and may simply be due to chance, particularly given (*i*) the closer relationship between theory and experiment with Type I and V rules, and (ii) the fact that each generation is a function of the previous generation, meaning that if one generation has an unusually low allele frequency, the subsequent generation likely will too. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

taking part and could earn up to a further $1 depending on their performance at the task.

Within each simulation, participants completed a single trial of a category-learning task in which they categorized 8 "amoebas" that varied along three binary dimensions into one of two categories; "good" and "bad" (there were always four amoebas in each category, see Fig. 6a). While somewhat artificial, such a task overlaps with many problems faced by real organisms, for instance categorizing threats in terms of whether to flee or freeze (Hébert, Versace, & Vallortigara, 2019), or how to categorizing novel objects into linguistic categories. The three dimensions in which amoebas varied were color (blue or orange), nucleus color (green or purple) and the presence/absence of spots. Participants were first shown the correct categorization of the amoebas for 15 s. They were then presented with all 8 amoebas in a random order and asked to categorize each one (pressing up for good and down for bad). In each simulation the true categorization was determined by one of three kinds of rule (Shepard, Hovland, & Jenkins, 1961) (Fig. 6a): Type I rules are based only on a single dimension (e.g. "blue amoeba are good"); Type II rules depend on the conjunction of two dimensions (e.g., "blue amoeba with spots and orange amoeba without spots are good"); Type V rules depend on all three dimensions and mimic a Type I rule, but with an exception (e.g. "amoeba with purple nuclei are good unless they are blue and do not have spots, in that case amoeba with green nuclei are good"). Type III, IV and VI rules also exist (Shepard et al., 1961), but were not included in our experiment. Human participants readily learn Type I rules, but find Type II rules harder and Type V rules harder still, with the exceptions being particularly difficult (Shepard et al., 1961). Of the 18 simulations, three were practice simulations (one with each type of rule), three used a Type I rule, six used a Type II rule and six used a Type V rule. Participants took part in the practice simulations first and then took part in all other simulations in a random order. Any participants who averaged < 7/8 correct categorizations across simulations using a Type I rule were assumed to have not been paying attention and were replaced by another participant (460 participants failed this check, roughly 16% of participants who completed the experiment).

As in the models presented above, within each simulation participants were assigned a simulated genome with 8 unlinked genes, each of which corresponded to one of the amoeba and had two alleles: *plastic* and *fixed*. The alleles did not affect participants' experience; regardless of their genome they were shown the correct categorization for 15 s before then being asked to categorize each of the 8 amoebas. The alleles did affect participant fitness, however; fitness was a function of the number of amoeba that were *either* categorized correctly *or* associated with a fixed allele:

$$F_i = max\left(\left(\frac{N}{4} - 1\right)^3, 0.0001\right), \tag{14}$$

where N is the number of amoebas that were *either* categorized correctly *or* associated with a fixed allele. As such, a participant with *fixed* alleles at all 8 loci would achieve maximal fitness even if they incorrectly categorized all 8 amoebas. In this way, *fixed* alleles do not rely on plasticity (i.e. human learning) to produce fitness benefits. This function is of a different form to the fitness functions used in the above models. It results in negligible fitness when $N < 5$, but rapid fitness gains for greater values. This was done to increase the strength of selection, given that even random behavior would predict $N \geq 4$, and preliminary theory suggested that such changes were required for robust results over the smaller population size and shorter timespan we used in the experimental simulations.

Once fitness was calculated for all participants in a generation, the next generation of participants was recruited. They inherited their genes via "sexual reproduction" among participants in the prior generation. Each incoming participant's two parents were chosen from the previous generation weighted by fitness and each gene was inherited from a randomly selected parent. Sexual reproduction was used as it accelerates the evolutionary dynamics (see supplementary material, section 2) and so makes equilibria more apparent over experimental timescales. Genes had a 5% chance of mutation, which produced the *plastic* allele with probability 0.75. In the first generation, participants' genes were initialized at random through mutation.

Based on the above theory, we predicted that (1) the overall frequency of the *fixed* allele would be highest in simulations with a Type V rule, and lowest in simulations with a Type I rule, and (2) in simulations with a Type V rule the frequency of the fixed allele would be higher for the exception amoebas than for the non-exception amoebas.

Data were analyzed with Bayesian Generalized Linear Mixed Models (GLMMs) using Markov Chain Monte Carlo (MCMC) methods to generate parameter estimates via the R package rjags (Plummer, Stukalov, & Denwood, 2016). All estimated values are based on an effective sample size of 3000, drawn from 3 parallel chains using the Gelman-Rubin statistic to ensure convergence (upper limit $\leq$ 1.01). Quoted estimates are the median sample and 95% central credible interval.

We modelled whether participants correctly categorized each amoeba (N = 288,000) as a Bernoulli variable. The probability of success was a function of the type of learning rule, an effect of the amoeba being an exception amoeba (only applicable for Type V rules) and a random effect for individual participants. In summary:

$$\text{Correct}_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = \text{rule\_effect}_r + \text{exception\_effect} * \text{exception}_i + \text{participant\_effect}_p$$

$$\text{rule\_effect} \sim \text{Normal}(0, 1000)$$

$$\text{exception\_effect} \sim \text{Normal}(0, 1000)$$

$$\text{participant\_effect} \sim \text{Normal}(0, \text{tau\_participant})$$

$$\text{tau\_participant} \sim \text{Gamma}(0.001, 0.001)$$

where r is the type of rule, *exception* is 1 if the amoeba was an exception amoeba (otherwise 0) and p is the numerical id of the participant.

We modelled whether each gene in the final generation of all simulations (N = 7200, excluding practice simulations) with an identical model to the above, except (1) the response variable was whether each gene was a fixed allele, as opposed to whether a participant's decision was correct, and (2) we included random effects for each simulation, instead of each participant.

3.2. Results

We found that Type I rules were easiest to learn (99% correct [99%, 99%] unaided by genes), Type II rules were slightly harder (89% [88%, 89%]), and Type V rules were harder still (85% [84%, 85%]). As predicted, the *fixed* allele was most frequent with a Type V rule (60% [59%, 63%]), less so with Type II (42% [40%, 44%]), and least frequent with Type I (28% [25%, 30%], Fig. 6b).

With a Type V rule the exception amoebas were particularly hard to learn (non-exception accuracy: 88% [87%, 88%], exception accuracy: 73%, [72%, 74%]). As predicted, the *fixed* allele was most frequent at loci corresponding to exception amoeba (76% [73%, 79%]) than at other loci (55% [53%, 57%], Fig. 6c).

In summary, the experimental results support our theoretical prediction that selection will favor the genetic fixation of traits that are difficult to acquire through plasticity. In the experiment, genetic fixation was more widespread when participants were presented with concepts that they found more difficult to learn (such as Type V rules). Moreover, genetic fixation was able to specifically target hard-to-identify items, such as the exception amoeba.

4. Discussion

Developmental plasticity has been argued to be an important part of the evolutionary process explaining how selection (i) can solve a *needle-in-a-haystack* problem (Hinton & Nowlan, 1987), (ii) can favor traits that rely on coordination across multiple individuals, specifically language (Pinker & Bloom, 1990), and (iii) can favor complex cognition (Deacon, 1997; Dennett, 2003; Morgan, 2016). Here, we built on these results using theoretical and experimental simulations to explore the impact of plasticity on genetic change. We find that natural selection preferentially targets genes that influence aspects of a trait that (i) other traits depend upon, (ii) that are associated with greater fitness benefits and (iii) that are less reliably acquired through plasticity. These results are mutually reinforcing; dependence can be imagined in terms of fitness benefits (the benefit of a trait might depend on other traits being present) or in terms of the probability of acquisition (the probability of acquiring a trait through plasticity might depend on other traits being present) (Morgan & Griffiths, 2015) and so our results concerning dependence would predict those concerning fitness benefits or the probability of acquisition.

In all cases, and in keeping with other work on this topic (Ancel, 1999, 2000; Hinton & Nowlan, 1987), natural selection does not extinguish plasticity as it is maintained by mutation despite negative selection (plasticity is extinguished in (Fontanari & Santos, 2017), though this is in the absence of mutation). However, the key result is that the developmental process becomes highly reliable. The *fixed* allele becomes widespread with regards to traits that are valuable or that are otherwise hard to acquire, while plasticity is focused on aspects of a trait that are readily acquired or relatively unimportant. In the context of real traits this is exactly what is observed in the cases of archer fish insect catching or bird song. In both cases plasticity remains—both behaviors are learnt (Feher et al., 2014; Schuster, 2007)—but learning is almost invisible. For example, a single observation of an adult archer fish is enough to markedly increase the performance of juveniles (Schuster et al., 2006), meanwhile even playback of a juvenile zebra finch's own malformed song is sufficient to prompt the development of species-typical song (Feher et al., 2014). This almost-hidden plasticity is precisely what our models suggest would be the leftovers of the Baldwin Effect.

These results are also consistent with studies of human language acquisition, where young children readily acquire language, despite the fact that their observations of adult speech seem to offer only a sample of the overall language (Chomsky, 1980). As above, such unexpectedly rapid and reliable language learning is consistent with past interactions between plasticity and genetic evolution and so our results support a role for the Baldwin Effect in language evolution (Deacon, 1997; Pinker & Bloom, 1990).

The effect of cultural inheritance is to oppose genetic change, with the prevalence of the fixed allele decreasing as the efficacy of cultural transmission ($p_{\max}$) increases. This highlights that genes are not the only means by which information can made heritable. In these instances, the traits are reliably acquired, not due to genetic influence, but instead because culture increases the efficacy of plasticity. This is, in effect, a “cultural Baldwin Effect”, with individual, plastic responses becoming highly heritable through cultural, rather than genetic, effects. This is particularly clear when there is a value or region of ρ (the frequency of the trait in the population) at which the probability of acquisition via plasticity increases rapidly (e.g., logistic or Heaviside functions when p_{min} is low and p_{max} is high). In these cases, because plasticity is initially an ineffective means to acquire the traits, a genetic capacity to acquire the traits without plasticity evolves. However, this genetic change causes the frequency of the trait to steadily increase and, should it increase enough, it will cause a rapid increase in the efficacy of plasticity. This will further increase ρ, but it will also weaken selection favoring the fixed allele, thereby causing the initial genetic response to be eroded by drift. In effect, the populations pass through a transitional period of genetic inheritance before reaching an equilibrium dominated by cultural inheritance.

As has been proposed elsewhere (Pinker & Bloom, 1990), this suggests that, because language is inherited culturally, there may be less need for a genetic basis to language. Moreover, variation between languages may prevent any genetic adaptations to specific linguistic features (Chater et al., 2009). This does not mean, however, that we should expect no genetic adaptations supporting a trait like human language. As our models show, even where a trait is extensively supported by cultural inheritance the frequency of the *fixed* allele remains above the level due to mutation alone (0.1), so even though plasticity can be highly effective, it nonetheless has limitations, and so selection will still seek out small benefits from genetic change. A solution, therefore, might be that human language is supported by indirect

genetic changes that increase our ability for high fidelity social transmission, prosocial interactions or tolerance of other individuals (Heyes, 2018; Laland, 2017). Changes like these may have provided enough learning opportunities to allow the otherwise-cultural evolution and inheritance of language.

An important assumption of the models presented here is that *fixed* and *plastic* are the only two possible alleles (a design choice made following existing work: (Fontanari & Santos, 2017; Hinton & Nowlan, 1987; Santos et al., 2015)). As such, future work may wish to include a wider range of alleles that have variable effects upon both the plasticity and success of an organism. In addition, it may be beneficial to consider a trait that is continuously varying: the traits we considered can be described as a binary string with each element being either present or absent. An alternative would be to have a continuous trait with an optimal value, with nearby values associated with modest payoffs (e.g. (Ancel, 1999, 2000; Lande, 2009; Scheiner et al., 2017)). We also assumed that mutation is more likely to produce a *plastic* allele than a *fixed* allele (a choice we made simply to make the increase in the *fixed* allele due to selection more apparent). While it seems plausible that the appearance of an allele that guarantees the successful acquisition of a fitness relevant trait (i.e., the *fixed* allele) is rarer than an allele that doesn't determine the phenotype in this way, it is less clear that mutations that permit or increase plasticity should be much more common that those that have deleterious effects on the phenotype. Nonetheless, we include supplementary results (see supplementary material, section 1) where mutation is assumed to be unbiased and the results were qualitatively unchanged.

Another important limitation of the results presented here is that none of the models include environmental change (either spatial or temporal), an omission that might seem unusual as environmental change has often been seen as central to the origin and maintenance of plasticity (Ancel, 1999; Badyaev, 2009; Baldwin, 1896; Scheiner et al., 2017). Nonetheless, we did not include it here as it goes beyond the scope of this work, whose focus is on how selection preferentially targets fixation. Further work may yet consider the addition of environmental change to the parameters we have already examined. Our expectation is that environmental change will not qualitatively change results in the absence of culture (unless environments are sufficiently unstable that an individual encounters highly different conditions across its lifetime) though it may promote the evolution of the fixed allele to greater frequencies (Ancel, 1999). With culture, the inclusion of environmental change may have a more significant impact on the results because cultural inheritance facilitates the persistence of information across time and space and so increases the odds that information encounters bouts of environmental change. Environmental change has been a significant focus of work in the cultural evolutionary literature where a common pattern is that very rapid temporal change tends to reduce the benefit of cultural learning, while spatial change and more moderate temporal change increase it (Aoki & Feldman, 2014; Boyd, Richerson, & Henrich, 2011; Ehn & Laland, 2012; Rendell, Fogarty, & Laland, 2010; Richerson & Boyd, 2013; Rogers, 1988).

The details of how plasticity affects the evolutionary process are yet to be fully understood. The results presented here build on existing work by using an abstract model to examine how plasticity affects the genetic evolution of a trait underpinned by multiple genetic loci. Our results suggest that plasticity shapes natural selection in a number of ways, targeting genetic change to maximize fitness without plasticity itself disappearing. The outcomes of our simulations match discussed cases of animal behavior, including human language, where the behavior remains plastic but is extremely reliably acquired. This suggests that these traits have evolved via an interaction between genetic evolution and plasticity, with plasticity shaping selection to favor genetic variants that support aspects of the traits that are difficult to learn. The inclusion of culture in this process has varied effects—reducing genetic change where culture provides an effective alternative to genetic inheritance, but increasing it when culture makes the traits more important. Across all our simulations, however, the consistent finding is that genetic equilibria are partially determined by the nature of plasticity. This implies that the details of developmental processes, such as the psychological mechanisms of learning, are essential to understanding the evolutionary process.

Author contributions

TJHM and TLG conceived of the project. TJHM created the models. TJHM and JWS created and ran the experiments. TJHM analyzed the data. All authors wrote the manuscript.

Acknowledgements

A preliminary version of this work was presented at the Annual Conference of the Cognitive Science Society and appears in the proceedings (Morgan & Griffiths, 2015). This work was funded by NSF grants 1456709 and 1408652.

Appendix A. Supplementary data

Experimental data is freely available at https://github.com/thomasmorgan/cognition-2019-data. This includes csv files of all data tables from the experiment, R code to compile the tables into a format suitable for analysis, and pre-compiled data as an R object. doi:https://doi.org/10.1016/j.cognition.2019.104165

References

Adler, D. S., Wilkinson, K. N., Blockley, S., Mark, D. F., Pinhasi, R., Schmidt-Magee, B. A., ... Gasparian, B. (2014). Early Levallois technology and the Lower to Middle Paleolithic transition in the Southern Caucasus. *Science, 345*(6204), 1609–1613. https://doi.org/10.1126/science.1256484.

Ancel, L. W. (1999). A quantitative model of the Simpson-Baldwin effect. *Journal of Theoretical Biology, 196*(2), 197–209. https://doi.org/10.1006/jtbi.1998.0833.

Ancel, L. W. (2000). Undermining the Baldwin expediting effect: Does phenotypic plasticity accelerate evolution? *Theoretical Population Biology, 58*(4), 307–319. https://doi.org/10.1006/tpbi.2000.1484.

Aoki, K., & Feldman, M. W. (2014). Evolution of learning strategies in temporally and spatially variable environments: A review of theory. *Theoretical Population Biology, 91*, 3–19. https://doi.org/10.1016/j.tpb.2013.10.004.

Badyaev, A. V. (2009). Evolutionary significance of phenotypic accommodation in novel environments: An empirical test of the Baldwin effect. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 364*(1520), 1125–1141. https://doi.org/10.1098/rstb.2008.0285.

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

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences, 108*(Supplement_2), 10918–10925. https://doi.org/10.1073/pnas.1100290108.

Chater, N., Reali, F., & Christiansen, M. H. (2009). Restrictions on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences, 106*(4), 1015–1020. https://doi.org/10.1073/pnas.0807191106.

Chomsky, N. (1980). On cognitive structures and their development: A reply to Piaget. In M. Piattelli-Palmarini (Ed.). *Language and learning: The debate between Jean Piaget and Noam Chomsky*. Harvard University Press.

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain.* New York, NY: W. W. Norton.

Dennett, D. (2003). The Baldwin effect: A crane, not a skyhook. In B. H. Weber, & D. J. Depew (Eds.). *Evolution and learning: The Baldwin effect reconsidered.* MIT press.

Dewitt, T. J., Sih, A., & Wilson, D. S. (1998). Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution, 13*(2), 77–81.

Earl, J. W., & McCleary, B. V. (1994). Mystery of the poisoned expedition. *Nature, 368*(6473), 683–684. https://doi.org/10.1038/368683a0.

Ehn, M., & Laland, K. (2012). Adaptive strategies for cumulative cultural learning. *Journal of Theoretical Biology, 301*, 103–111. https://doi.org/10.1016/j.jtbi.2012.02.004.

Feher, O., Suzuki, K., Okanoya, K., Ljubicic, I., & Tchernichovski, O. (2014). Birds tutored with their own developing song produce wildtype-like song as adults. *Proceedings of the 10th International Conference on the Evolution of Language.* (pp. 433–434).

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics, 78*, 737–756.

Fontanari, J. F., & Santos, M. (2017). The revival of the Baldwin effect. *European Physical Journal B, 90*(10), https://doi.org/10.1140/epjb/e2017-80409-8.

Gabora, L. (2008). The cultural evolution of socially situated cognition. *Cognitive Systems Research, 9*(1–2), 104–114. https://doi.org/10.1016/j.cogsys.2007.05.004.

Gerullis, P., & Schuster, S. (2014). Archerfish actively control the hydrodynamics of their jets. *Current Biology, 24*(18), 2156–2160. https://doi.org/10.1016/j.cub.2014.07.059.

Hébert, M., Versace, E., & Vallortigara, G. (2019). Inexperienced preys know when to flee or to freeze in front of a threat. *PNAS,* 1–3. https://doi.org/10.1073/pnas.1915504116.

Heyes, C. (2018). *Cognitive gadgets.* Cambridge, MA: Harvard University Press.

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

Hogan, J. (1994). Structure and development of behavior systems. *Psychonomic Bulletin & Review, 1*(4), 439–450. https://doi.org/10.3758/BF03210948.

Laland, K. N. (2017). *Darwin's unfinished symphony.* Princeton University Press.

Laland, K. N., Wray, G. A., & Hoekstra, H. E. (2014). Does evolutionary theory need a rethink? *Nature,* 4–7.

Lande, R. (2009). Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology, 22*(7), 1435–1446. https://doi.org/10.1111/j.1420-9101.2009.01754.x.

Lloyd Morgan, C. (1896). On modification and descent. *Science, 4*(99), 733–739.

Lycett, S. J., & Von Cramon-Taubadel, N. (2013). A 3D morphometric analysis of surface geometry in Levallois cores: Patterns of stability and variability across regions and their implications. *Journal of Archaeological Science, 40*(3), 1508–1517. https://doi.org/10.1016/j.jas.2012.11.005.

Morgan, T. J. H. (2016). Testing the cognitive and cultural niche theories of human evolution. *Current Anthropology, 57*(3), 370–377. https://doi.org/10.1086/686531.

Morgan, T. J. H., & Griffiths, T. L. (2015). What the Baldwin effect affects. *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 1643–1648). .

Morgan, T. J. H., & Harris, P. L. (2015). James Mark Baldwin and contemporary theories of culture and evolution. *European Journal of Developmental Psychology, 5629*(November), 1–12. https://doi.org/10.1080/17405629.2015.1074068.

Muller, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research, Fundamental and Molecular Mechanisms of Mutagenesis, 1*(1), 2–9. https://doi.org/10.1016/0027-5107(64)90047-8.

Osborne, H. F. (1896). A mode of evolution requiring neither natural selection nor the inheritance of acquired characteristics. *Transactions of the New York Academy of Sciences, 15*, 141–148.

Pigliucci, M., Murren, C. J., & Schlichting, C. D. (2006). Phenotypic plasticity and evolution by genetic assimilation. *The Journal of Experimental Biology, 209*(Pt 12), 2362–2367. https://doi.org/10.1242/jeb.02070.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*, 707–784.

Plummer, M., Stukalov, A., & Denwood, M. (2016). *rjags: Bayesian graphical models using MCMC*.

Rendell, L., Fogarty, L., & Laland, K. N. (2010). Rogers' paradox recast and resolved: Population structure and the evolution of social learning strategies. *Evolution, 64*(2), 534–548. https://doi.org/10.1111/j.1558-5646.2009.00817.x.

Richards, R. J. (1987). *Darwin and the emergence of evolutionary theories of mind and behavior.* Chicago/London: University Of Chicago Press.

Richerson, P. J., & Boyd, R. (2013). Rethinking paleoanthropology: A world queerer than we supposed. In G. Hatfield, & H. Pittman (Eds.). *Evolution of mind, brain, and culture* (pp. 263–302). University of Pennsylvania Press.

Rogers, A. R. (1988). Does biology constrain culture. *American Anthropologist, 90*(4), 819–831.

Santos, M., Szathmáry, E., & 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. https://doi.org/10.1016/j.jtbi.2015.02.012.

Scheiner, S. M. (1993). Genetics and evolution of phenotypic plasticity. *Annual Review of Ecology and Systematics, 24*, 35–68.

Scheiner, S. M., Barfield, M., & Holt, R. D. (2017). The genetics of phenotypic plasticity. XV. Genetic assimilation, the Baldwin effect, and evolutionary rescue. *Ecology and Evolution,* (April), 8788–8803. https://doi.org/10.1002/ece3.3429.

Schuster, S. (2007). Archerfish. *Current Biology, 17*(13), 494–495. https://doi.org/10.1016/j.cub.2007.04.014.

Schuster, S., Wöhl, S., Griebsch, M., & Klostermeier, I. (2006). Animal cognition: How archer fish learn to down rapidly moving targets. *Current Biology, 16*(4), 378–383. https://doi.org/10.1016/j.cub.2005.12.037.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied, 75*(13), https://doi.org/10.1017/CBO9781107415324.004.

Simpson, G. G. (1953). The Baldwin effect. *Evolution, 7*(2), 110–117.

Temple, S., Hart, N. S., Marshall, N. J., & Collin, S. P. (2010). A spitting image: Specializations in archerfish eyes for vision at the interface between air and water. *Proceedings of the Royal Society B: Biological Sciences, 277*(1694), 2607–2615. https://doi.org/10.1098/rspb.2010.0345.

Thorpe, W. H. (1961). *Bird song.* Cambridge, UK: Cambridge University Press.

Via, S., Gomulkiewicz, R., De Jong, G., Scheiner, S. M., Schlichting, C. D., & Van Tienderen, P. H. (1995). Adaptive phenotypic plasticity: Consensus and controversy. *Trends in Ecology & Evolution, 10*(5), 212–217. https://doi.org/10.1016/S0169-5347(00)89061-8.

Wagner, G. P. (2014). *Homology, genes, and evolutionary innovation.* Princeton University Press.

Weber, B., & Depew, D. (2003). *Evolution and learning: The Baldwin effect reconsidered.* MIT Press.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution.* Oxford University Press.