

The Cultural Origins of Cognitive Adaptations  
David Papineau

1 Introduction

According to an influential view in contemporary cognitive science, many human cognitive capacities are innate. The primary support for this view comes from ‘poverty of stimulus’ arguments. In general outline, such arguments contrast the meagre informational input to cognitive development with its rich informational output. Consider the ease with which humans acquire languages, become facile at attributing psychological states (‘folk psychology’), gain knowledge of biological kinds (‘folk biology’), or come to understand basic physical processes (‘folk physics’). In all these cases, the evidence available to a growing child is far too thin and noisy for it to be plausible that the underlying principles involved are derived from general learning mechanisms. This only alternative hypothesis seems to be that the child’s grasp of these principles is innate. (Cf. Laurence and Margolis, 2001.)

At the same time, it is often hard to understand how this kind of thing could be innate. How exactly did these putatively innate cognitive abilities evolve? The notion of innateness is much contested—we shall return to this issue at the end of the paper—but on any understanding the innateness of some complex trait will require a suite of genes which contributes significantly to its normal development. Yet, as I shall shortly explain, there are often good reasons for doubting that standard evolutionary processes could possibly have selected such suites of genes.

In this paper I want to outline a non-standard evolutionary process that could well have been responsible for the genetic evolution of many complex cognitive traits. This will in effect vindicate cognitive nativism against the charge of evolutionary implausibility. But at the same time it will cast cognitive nativism in a somewhat new light. The story I shall tell is one in which the ancestral learning of cognitive practices plays a crucial role, and in which this ancestry has left a mark on contemporary cognitive capacities, in a way that makes it doubtful that there is anything in them that is strictly ‘innate’, given a normal understanding of this term. For, if my account of the evolution is right, it seems likely that acquisition of information from the environment will always continue to be involved alongside genes in the ontogeny of such traits. On the picture I shall develop, then, we pay due respect to ‘poverty of the stimulus’ considerations—certainly the ease and reliability with which many cognitive powers are acquired shows that there are genes which have been selected specifically to facilitate these powers—but this does not mean that they are ‘innate’ in any stronger sense—for their acquisition will still depend crucially on information derived from environmental experience.

2 An Evolutionary Barrier

Why do I say that that standard evolutionary processes cannot account for the selection of the suites of genes behind complex cognitive traits? Cannot nativists simply offer the normal adaptationist explanation, and say that the relevant genes were selected because of the selective advantages they offered? However, there is a familiar difficulty facing such adaptationist accounts of complex traits, which we might call the ‘hammer and nail’ problem. If some phenotypic trait depends on a whole suite of genes, it is not enough for an adaptationist evolutionary explanation that the phenotype as a whole should be adaptive. After all, if the relevant genes originally arose by independent mutation, then the chance of their all occurring together in some individual would have been insignificant, and even if they did co-occur, they would quickly have been split up by sexual reproduction. So the fact that they would have yielded an advantage, if they had all co-occurred, is no explanation at all of how they all became common. Rather each gene on its own needs to bring some advantage, even in the absence of the other genes. It is by no means clear that this requirement will be satisfied for the paradigm examples of putatively innate cognitive powers. Is there any advantage to the ‘mind-reading’ folk psychological ability to tell when someone else can see something, if you don’t yet know how this will lead them to behave, or vice versa? Is there any advantage to being disposed to identify anaphoric linguistic constructions, if you don’t yet know that languages have a systematic way of marking subject-object position, or vice versa?<sup>111</sup> (Is there any advantage to a hammer, if there are no nails to hit with it, or any advantage to nails, if there is no hammer to hit them with?)

---

<sup>111</sup> The socially cooperative nature of language presents another kind of evolutionary hurdle: what if the use of one person having genes for language, if nobody else yet has them? In the interests of generalizing over non-cooperative cognitive capacities as well, I shall not stress this particular difficulty in what follows. However, the points made about social learning in section 6 indicate the obvious mechanism by which it could have been surmounted.

Notoriously, the major proponents of cognitive nativism have dealt with this challenge by largely ignoring it. Both Noam Chomsky and Jerry Fodor are famous for insisting that evolutionary considerations have no relevance to cognitive science. In their view, attempts to pin the down the evolutionary origin of cognitive traits are at best entertaining speculations, and at worst a distraction from serious empirical investigation (Chomsky, 1972, Fodor, 2000). However, this attitude simply fails to engage with the above challenge.<sup>2[2]</sup> Questions about evolutionary origins may be difficult, but this doesn't alter the fact that a posited suite of genes can't actually exist if they can't possibly have evolved.

In the last decade or so, the self-styled 'Evolutionary Psychology' movement has married the nativism of Chomsky and Fodor with a positive concern for evolutionary questions, suggesting that a greatly expanded range of cognitive 'modules' (including modules for cheater-detection, mate-selection, and so on, as well as for language and the folk theories mentioned above) are evolutionary adaptations produced by selective pressures operating in the 'Environment of Evolutionary Adaptation' (Barkow, Cosmides and Tooby (eds), 1992). However, it cannot be said that the Evolutionary Psychology movement has properly engaged with the 'hammer and nail' issue. By and large, its adherents have been content to adopt a simple 'adaptationist' stance, assuming from the start that natural selection has the power bring about adaptive traits when they are needed. There is little in the writings of committed Evolutionary Psychologists to assuage the doubts of sceptics who feel that the selective barriers faced by innate cognitive modules are reason to doubt that such innate modules exist. (However, see Pinker and Bloom, 1990, esp. section 5.2.)

### 3 Learning as a Basis for Genetic Advantage

In this paper, I want to consider a possible mechanism which might explain how the evolution of complex cognitive abilities might overcome 'hammer and nail' hurdles. Such hurdles arise when a specific gene is only selectively advantageous given a context of pre-existing cognitive traits. I shall show that such a gene can nevertheless be selected even in the absence of other genes which fix the pre-existing traits. The central thought of this paper is that it will be enough for such selection if those other traits are being learned. After all, what is required is that the other pre-existing traits should be present, not that they be genetically fixed, and there is no obvious reason why learning should suffice for this.

The details of this suggestion will be examined at length in what follows. But I hope it will be immediately clear how it promises to overcome the 'hammer and nail' problem. Take some complex cognitive ability. As long as this ability is being learned, then this itself may create an environment in which genes that contribute elements of this ability will be selected. In effect, once the ability is being learned, then the relevant genes will start being selected precisely because they lighten the burden of learning.

This suggests the intriguing possibility that the innate modules so emphasized by recent nativist opinion are all 'fossilized' versions of abilities which originally arose from general learning mechanisms. If this right, then the genetic shaping of the modern human mind, far from demonstrating the impotence of general learning, is a testament to its fecundity.

I have introduced this suggestion by emphasizing the possibility of selective obstacles of the 'hammer and nail' variety. Some readers may remain unconvinced that this is a real problem. In particular, they may have felt I was too quick to dismiss the possibility that genes for the various components of complex cognitive traits might each be selectively advantageous on their own. Why shouldn't there be room for the strategy Richard Dawkins employs in Climbing Mount Improbable (1996), where he shows, against those who argue that a part of a wing is no advantage at all, say, just how even a part of a wing may be better than nothing? Similarly, despite first appearances, maybe there is some advantage to being able to tell whether another organism can see something, even without knowing

---

<sup>2[2]</sup> Sometimes Chomsky and Fodor suggest that our innate linguistic powers may not be adaptations after all, but simply 'spandrel'-like by-products of other evolutionary developments (Chomsky, 1988, Fodor, 2000). This could be read as an implicit recognition of the 'hammer and nail' problem facing any simple adaptationist story. Still, the idea that all our innate linguistic powers are spandrels is difficult to take seriously. If a simple adaptationist account is ruled out, a far more plausible alternative is a complex adaptationist account, not a miracle. In effect, this is what I offer below.

what this will make them do . . . (Maybe hammers would be useful, even without nails, for banging other things . . .)

I shall not take direct issue with this response. For what it is worth, I suspect that ‘hammer and nail’ obstacles are common enough in cognitive evolution, and that many of the cognitive traits that interest us simply could not have evolved with the help of prior stages when they were learned. But I do not need to defend this strong claim here. This is because the selective process I shall focus on does not require the absolute impossibility of evolving hammers without nails. Maybe many of the elements in the human understanding of mind, say, are of some biological advantage on their own, and maybe this alone could have led to the independent selection of genes which variously fix these elements. It is consistent with this that each of these elements are much more advantageous when found in conjunction with the rest of the understanding of mind, and thus that the initial selection of the relevant genes would have proceeded all the faster in contexts where other parts of understanding of mind was already being acquired from general learning processes. This argues that the kind of selection pressures I shall be exploring would have played a significant role whenever learning helped to foster complex cognitive structures, including cases when there was no absolute ‘hammer and nail’ obstacle to the selection of genes for those structures in the absence of learning. Given this, even readers who feel that I have overstated the ‘hammer and nail’ issue should still find what follows of interest.

#### 4 Genetic Takeovers

Let me now give a more detailed analysis of the basic selective process I am interested in. It will be helpful in this connection to turn away from human cognition for a while and consider a simple example of bird behaviour. The woodpecker finches of the Galapagos Islands use twigs or cactus spines to probe for grubs in tree branches (Tebich et al. 2001; see also Bateson 2004). This behaviour involves a number of component dispositions—finding possible tools, fashioning them if necessary, grasping them in the beak, using them to probe at appropriate sites. As it happens, the overall grub-seeking behaviour of the finches displays a high degree of innateness (though see section 14 below). Yet the evolution of this innateness would seem to face a severe version of the ‘hammer and nail’ obstacle. None of the component dispositions is of any use by itself. For example, there is no advantage in grasping tools if you aren’t disposed to probe with them, and no advantage to being disposed to probe with tools if you never grasp them. This makes it very hard to see how genes for the overall behaviour could possibly have been selected for. In order for the behaviour to be advantageous, all the components have to be in place. But presumably the various different components are controlled by different genes. So any biological pay-off would seem to require that all these genes be present together. However, if these genes are initially rare, it would be astronomically unlikely that they would ever co-occur in one individual, and they would quickly be split up by sexual reproduction even if they did. So the relevant genes, taken singly, would seem to have no selective advantage that would enable them to be favoured by natural selection.

However, now suppose that, before the grub-seeking behaviour became innate in the finches, there was a period where the finches learned to catch grubs, by courtesy of their general learning mechanisms. This could well have itself created an environment where each of the genes that facilitate the overall behaviour would have been advantageous. For each of these genes, on its own, would then have the effect of fixing one component of the grub-seeking behaviour, while leaving the other components to be acquired from learning. And this could itself have been advantageous, in reducing the cost and increasing the reliability with which the overall behaviour was acquired. The result would then be that each of the genes would be selected for, with the overall behaviour thus coming increasingly under genetic control. (There is a general issue here, to do with the relative selective advantages of genes and learning, which I shall address in the next section. For the moment let us simply suppose that the advantages due to genes, such as increased speed and reliability of acquisition, are not outweighed by any compensating disadvantages, such as reduced ontogenetic plasticity.)

Here is a general model of this kind of process, which I shall call ‘genetic takeover’.<sup>3[3]</sup> Suppose  $n$  subtraits,  $P_i$ ,  $i = 1, \dots, n$ , are individually necessary and jointly sufficient for some adaptive phenotype  $P$ , and that each subtrait is no good without the others. (Thus: finding tool materials, fashioning them,

---

<sup>3[3]</sup> A common alternative term for this process is ‘genetic assimilation’ (cf. Hinton and Nowlan, 1987, Turney et al., 1996, Avital and Jablonka 2001, Godfrey-Smith 2003, Papineau, 2005). However, this term was originally coined by C.H. Waddington (1953, 1957, 1961), and there is some controversy as to whether he had the same process in mind (Bateson 2004, Griffiths 2006, Papineau, 2006). ‘Genetic takeover’ avoids this exegetical issue.

grasping them, . . .) Suppose further that each sub-trait can either be genetically fixed or acquired through learning, with alternative alleles at some genetic locus either genetically determining the sub-trait or leaving it plastic and so available for learning. So, for sub-trait  $P_i$ , we have allele  $G_i$  which genetically fixes  $P_i$ , and allele(s)  $L_i$  which allows it to be learned.

To start with, the  $G_i$ s that genetically determine the various  $P_i$ s are rare, so that it is highly unlikely that any individual will have all  $n$   $P_i$ s genetically fixed. Still, having some  $P_i$  genetically fixed will reduce the amount of learning required to learn the overall behaviour. (If you are already genetically disposed to grab suitable twigs if you see them, you will have less to do to learn the rest of the tool-using behaviour.) Organisms with some  $G_i$ s will thus have a head start in the learning race, so to speak, and so will be more likely to acquire the overall phenotype. So the  $G_i$ s that give them the head start will have a selective advantage over the  $L_i$ s. Natural selection will thus favour the  $G_i$ s over the  $L_i$ s, and in due course will drive the  $G_i$ s to fixity.<sup>4[4]</sup>

This genetic takeover model is a simplification of one developed by Hinton and Nowlan (1987). They ran a computer simulation using a ‘sexually reproducing’ population of neural nets, with an ‘advantageous phenotype’ that required the 20 connections in their neural nets all to be set at ‘1’ rather than ‘0’. Insofar as it was left to solely to ‘genes’ and sexual sorting, there was a miniscule chance of hitting the advantageous phenotype, and so genes for ‘1’s were not selected. However, once the nets could ‘learn’ during their individual lifetimes to set their connections at ‘1’, then this gave genes for ‘1’s an advantage (since they increased the chance of so learning the advantageous overall phenotype), and in this context these genes then progressively replaced the alternative alleles which left the connections to learning.

It is worth spelling out exactly how the genetic takeover model offers a way of overcoming selective ‘hammer and nail’ obstacles. At first it may seem that each  $G_i$  will have no selective advantage on its own, given that it only fixes one  $P_i$ , which isn’t of any use without the other  $P_i$ s. But in a context where the various  $P_i$ s can also be learned, each  $G_i$  does have a selective advantage on its own, even in the absence of the other  $G_i$ s, precisely because it makes it easier to learn the rest of  $P$ . Even in the absence of other  $G_i$ s at other loci, any given  $G_i$  will still be favoured by natural selection, because it will reduce the learning load and so make it more likely that its possessor will end up with the advantageous phenotype  $P$ . This is what drives the progressive selection of the  $G_i$ s in the model. Each  $G_i$  is advantageous whether or not there are  $G_i$ s at other loci, simply because having a  $G_i$  rather than an  $L_i$  at any given locus will reduce the amount of further learning needed to get the overall  $P$ .

Much previous discussion of this kind of model has taken place under the heading of the ‘Baldwin Effect’. This notion traces back to James Mark Baldwin (1896) and others evolutionary theorists at the end of the nineteenth century. While it is not always clear what these thinkers originally had in mind, the ‘Baldwin Effect’ is now standardly understood to refer to any selective process whereby some trait  $P$  is brought under genetic control as a result of previously being under environmental control. At first pass, of course, the Baldwin Effect sounds like Lamarckism, and indeed many commentators have argued that there can be no legitimate Darwinian mechanism fitting the specifications of the Baldwin Effect. (How can the prior environmental control of  $P$  possibly matter to selection, given that those who benefit from environmentally acquiring some trait won’t pass on any genes for that trait to their offspring? Cf. Watkins, 1999.)

In this paper I shall generally steer clear of the intricate literature on the Baldwin Effect. But, for what it is worth, the genetic takeover model does at least provide one legitimate way in which a trait can come under genetic control as a result of previously being under environmental control. In this model the population of organisms moves from a stage in which the overall  $P$  is initially acquired by learning to a stage where it is genetically fixed. Moreover, the first stage is essential to the second, in that the alleles  $G_i$  which together genetically fix  $P$  would have had no initial selective advantage were  $P$  not previously learned.

---

<sup>4[4]</sup> This model should be handled with care. There is no need to think of the relevant loci as somehow ‘dedicated’ to the related phenotypes—the idea is only that each may be occupied by an allele  $G_i$  which (produces a protein) that causes the phenotype  $P_i$  in question; the alternative allele(s)  $L_i$  needn’t be thought of as somehow specifically ensuring that  $P_i$  is learnable, as opposed to simply doing nothing to stop  $P_i$  being one of the many phenotypes that can be acquired from general learning mechanisms. Relatedly, it is only for purposes of expository simplification that I assume that the  $G_i$  alleles on their own determine recognizable phenotypic components  $P_i$ ; what is crucial is solely that the  $G_i$ s determine proteins that somehow make learning the overall  $P$  easier. I shall return to this point in my final section.

## 5 Genes versus Learning

Let me now address the question of the relative benefits of learning and genetic control. In the last section I took it for granted that genetic takeover will generally be selectively advantageous. That is, I supposed that the  $L_i$  alleles which leave some element of an adaptive phenotype to learning will in general be outcompeted by the  $G_i$ s which ensure that those components become genetically fixed. However, it is by no means automatic that this should be so. There are costs as well as benefits to genetic control, and genetic takeover therefore requires that the latter outweigh the former.

Let me begin by detailing the possible advantages of genetic takeover. At first sight it may be unclear why there should be any such advantages. If the relevant phenotype will be acquired by learning in any case, as in our cases of possible genetic takeover, what extra advantage will derive from genetic determination? The immediate answer is that the relevant phenotype won't always be acquired in any case, if it is not genetically fixed. Learning is hostage to the quirks of individual history, and a given individual may fail to experience the environments required to instil some learned trait. Moreover, even if the relevant environments are reliably available, the business of learning some phenotype may itself involve immediate biological costs, delaying the time at which it becomes available, and diverting resources from other activities. In particular, the fact that the phenotype needs to be learned, rather than coming for free with the genome, may mean that organisms are limited in their opportunities to learn further adaptive traits, and are thus biologically disadvantaged for this reason.<sup>5[5]</sup>

On the other side must be placed the loss of flexibility that genetic fixity may entail. Learning will normally be adaptive across a range of environments, in each case producing a phenotype that is advantageous in that specific environment. By contrast, genes which fix traits that are only adaptive in some given environment will be of no biological advantage if the environment changes so as to render that trait maladaptive. In circumstances of environmental variability, an organism with genes that fix some trait may thus be less fit than one which relies on learning to tailor its phenotype to its environment.<sup>6[6]</sup>

As a general rule, then, we can expect that genetic fixity will be favoured when there is long-term environmental stability, and that learning will be selected for when there are variable environments. Given environmental stability, genetic fixity will have the aforementioned advantages of reliable and cheap acquisition. But these advantages can easily be outweighed by loss of flexibility when there is significant environmental instability. Exactly how the pluses and minuses of genetic control versus learning work out will depend on the parameters of particular cases.<sup>7[7]</sup> For the moment, I shall continue to assume that we are dealing with cases where genetic control has the overall biological advantage. I shall have more to say about this issue in section 12 below.

## 6 The Significance of Social Learning

It may seem that my hypothesized mechanism for circumventing hammer-and-nail obstacles simply trades in one kind of improbability for another, substituting improbabilities of complex learning for improbabilities of genetic co-occurrence. I have focused on cases where some complex adaptive phenotype  $P$  consists of various sub-parts  $P_i$ , none of which are adaptive on their own. And I have answered the puzzle of how genes for these  $P_i$ s could be selected, if none is advantageous on its own, by suggesting that these genes will become advantageous if the overall  $P$  can be learned. However, if

---

<sup>5[5]</sup> A side-effect of genetic control is thus the 'assimilate-stretch' process emphasized in Avital and Jablonka (2001): once some cognitive capacity is taken under genetic control and learning resources are thereby freed up, then organisms gain the opportunity to learn more sophisticated elaborations of that capacity, which may in turn be taken under genetic control, . . . and so on.

<sup>6[6]</sup> The loss of flexibility due to increased genetic control may well extend beyond the specific phenotype that is taken over genetically. When some trait that is originally shaped by some suite of relatively general learning mechanisms comes under genetic control, this may not be a simple matter of that trait alone being switched, so to speak, from the control of those general learning mechanisms to direct genetic control. For it is possible that the general learning repertoire will itself be affected by such switching. Perhaps bringing one trait under genetic control can make an organism less efficient at learning other traits. For example, if you are genetically predisposed towards folk psychology, then perhaps this will limit your ability to learn about non-psychological mechanisms. Commentators are somewhat divided on how far this danger is real (cf. Godfrey-Smith, 2003, Bateson, 2004).

<sup>7[7]</sup> For a detailed quantitative analysis of the relative costs of learning and genetic control, see Mayley (1996). Note that, in contexts where learning has the biological advantage over genetic fixity, then we might well find 'reverse Baldwin effects', where some trait originally under genetic control comes to depend on learning instead.

the overall P is complex, and none of its parts advantageous on their own, won't there equally be a problem about learning all of P?

Consider our Galapagos finches once more. The P<sub>i</sub>s there were finding tools, fashioning them, grasping them, using them to probe . . . Now just as there was no reproductive advantage in finding tools, or fashioning them, if you don't know how to grab them, or probe, and vice versa, neither will there be any psychological reward in having any of these dispositions without the others. However, this is likely to block the individual learning of the various dispositions, since such learning hinges on psychological reward, and it is extremely unlikely that random behaviour generation will ever lead some animal to perform all the requisite actions in sequence. Maybe the improbabilities involved in learning won't be as bad as those operating at the genetic level. But they may still be bad enough to ensure that, even after you have one gene G<sub>i</sub> for one of the P<sub>i</sub>s, there is no real chance of learning the rest of P, and so no real selective pressure in favour of that G<sub>i</sub>. So we still seem to face a 'hammer and nail' problem even after we introduce the possibility of learning, and for the same reason—the component P<sub>i</sub>s don't bring any pay-off on their own.

However, suppose now that we are dealing with organisms that are capable of social as well as individual learning. Maybe there is a very low probability of any individual with some one G<sub>i</sub> acquiring all the further elements of P via individual trial-and-error learning. But now suppose that the relevant population of animals has a culture of doing P—imagine, say, that the ancestors of the present Galapagos finches acquired their tool-using behaviour, not from individual trial-and-error learning, but via social learning from other finches who were already displaying it. This could then radically reduce the improbability of learning the various elements of P, and so could serve to render the G<sub>i</sub>s advantageous after all. If there is a real chance of learning all the requisite elements of P from others, then as before each G<sub>i</sub> could be selected because it increased the speed and reliability with which P is learned.

It is interesting to note that, when social learning plays a role in this way, then the 'genetic takeover' of P will qualify as a 'Baldwin Effect' for a reason over and above that outlined in the last section. The requirements for a 'Baldwin Effect', recall, were that some trait P is brought under genetic control as a result of previously being under environmental control. When a genetic takeover of P is facilitated by social learning, then we have this requirement being satisfied for the reason that the relevant genes would not be selected without the prior culture of P. The relevant G<sub>i</sub>s have a selective advantage specifically because of the pre-existing socially learned culture—without the culture, it would be too hard for individuals to learn the further elements of P needed to render G<sub>i</sub>s advantageous. A gene which helped a finch to identify suitable twigs would have no biological virtue if the finch's only way of acquiring the rest of the tool-using behaviour was by individual trial-and-error learning. However, once these things can be learned by example from the other finches, then the gene becomes advantageous in a way it wasn't before. In short, the genes for P get selected as a result of P previously being socially learned.

This way of satisfying the Baldwin requirement is not the same as that described in the last section. There the idea was simply that each G<sub>i</sub> would get selected because it made it easier to learn the rest of P. There was no assumption there that this learning depended on some prior culture. Any kind of learning, even non-social trial-and-error learning, would ensure that the G<sub>i</sub>s moved towards fixity via intermediate stages where the components of P were learned—and this in itself, as I pointed out, would give us one kind of 'Baldwin Effect'. I have now added in the further thought that in many cases learning the components of P may only be possible because other animals are already displaying P as an exemplar for social learning—this gives us another way of satisfying the Baldwin requirement that the selection of genes for P depends on P previously being learned. (To see clearly that these ways are different, note that, if there were any cases where individual trial-and-error learning created selection pressures for the G<sub>i</sub>s in the absence of social learning, counter to this section's line of argument, then we would still get genetic takeover even in radically unsocial species where no individual ever observes P in another organism at all. Here we would have a Baldwin Effect in the first sense—the G<sub>i</sub>s will get driven to fixity via helping each organism to learn P individually—but not in the second sense—the genetic takeover doesn't depend on other animals already learning P and providing a model for learning.)

In what follows, I shall focus on cases where social learning does play a crucial role in facilitating genetic takeover, and thus where the Baldwin requirement is satisfied twice over. In itself, this double

satisfaction of the Baldwin requirement is merely a conceptual oddity. It is of no special theoretical significance that certain possible processes should fit the half-formed ideas of an unimportant nineteenth-century theorist in two different ways.<sup>8[8]</sup> However, there is independent reason to think that these doubly Baldwinian processes are biologically significant—they offer a plausible selective mechanism whereby complex cognitive adaptations can come under genetic control. To repeat the argument so far, it is often puzzling how complex cognitive abilities can be selected for, given that their various components seem of no biological components on their own. However, we have seen how such selection can indeed take place if the ability in question is initially learnable; moreover, we have seen how such learning can be rendered possible by cultural transmission, even in cases where it would be beyond the powers of individual trial-and-error learning. These points in themselves provide reason seriously to investigate the genetic takeover of culturally transmitted traits, quite apart from the fact that they satisfy Baldwin's requirements twice over.

## 7 Getting Cultures Started

In the last section I argued that social learning can facilitate behaviours that are beyond the reach of individual trial-and-error learning, and thus render those behaviours available for 'genetic takeover'. However, there are a number of complexities hidden under this simple appeal to 'social learning'.

For a start, there is an obvious worry that the appeal to social learning merely postpones the problem that many cognitive practices are too complex to be acquired by individual trial-and-error learning. After all, a culture has to get started somehow. There has to be some initial stage where the cognitive practice is introduced to the population, in order that individuals can start learning it from others who already display it. The only obvious way for this to happen is for some lucky or exceptional individual to strike on the practice by some individual means. However, this may seem to be in obvious tension with the idea that social learning helps precisely with practices that are too complex to be acquired by individual trial-and-error learning.

However, this tension is more apparent than real. Think of social learning as a process which takes us from one individual learning P to its becoming socially learnable by all. This can make it highly likely that P will become prevalent, even though it's very hard for any given individual to get P from trial-and-error. Suppose that the chance of any given individual learning P by trial-and-error is  $k$ , and that there are  $n$  individuals in the population. Then the probability of at least one individual arriving at P by trial-and-error will be  $1-(1-k)^n$ , and this can be high even if  $k$  is low. (For example, even if there is only a 10% chance of any given individual will get P from trial and error, it is 88% likely that at least one individual in a group of 20 will so get it.) In short, social learning switches the probability that any given individual X will somehow learn P, from the low (10%) probability that X will acquire P from individual trial-and-error, to the high (88%) probability that someone will acquire P from individual trial-and-error.

## 8 Varieties of Social Learning

Let us now look more closely at the idea of 'social learning' itself. My last section simply assumed that 'social learning' will ensure that any adaptive cognitive ability—any 'good trick', as Daniel Dennett terms it (1991)—will spread throughout a population as soon as any one member acquires it from trial-and-error learning. However, this cannot be taken for granted. There are different kinds of social learning, displayed by different species of animals, and by no means all of them will automatically transfer the kind of 'good tricks' at issue here from individual to population.

At its most general, 'social learning' refers to any processes by which the display of some behaviour by one member of a species increases the probability that other members will perform that behaviour. However, this covers a numbers of different mechanisms. We can usefully distinguish (cf. Shettleworth, 1998, Tomasello, 2000):

(i) Stimulus Enhancement. Here one animal's doing P merely increases the likelihood that other animals' behaviour will become conditioned to relevant stimuli via individual learning. For example,

---

<sup>8[8]</sup> Did Baldwin himself have my doubly Baldwinian process in mind? It is not clear. He did on occasion mention social learning as important for his topic, and later writers have also alluded specifically to social learning when discussing the Baldwin Effect (Baldwin, 1896, Watkins, 1999). But I have found no explicit analysis in the literature of why social learning matters in this context.

animals follow each other around—novices will thus be led by adepts to sites where certain behaviours are possible (pecking into milk bottles, say, or washing sand off potatoes) and so be more likely to acquire those behaviours by individual trial-and-error.

(ii) Goal Emulation. Here animals will learn from others that certain resources are available, and then use their own devices to achieve them. Thus they might learn from others that there are ants under stones, or berries in certain trees.

(iii) Blind Mimicry. Here animals copy the movements displayed by others, but without appreciating to what end these movements are a means.

(iv) Learning about Means to Ends. Here animals grasp that some conspecific's behaviour is a means to some end, and copy it because they want that end.

We can take it that the first two kinds of social learning will be present in a wide range of species. They require nothing more than a tendency for animals to move around together, plus powers of instrumental learning (i), or pre-existing abilities to exploit resources once they are detected (ii). Blind mimicry (iii) is less common: while it is possible that some non-human animals have this capacity, it is by no means universal, even in mammals and birds (Shettleworth, 1998). Full-blooded appreciation of the relevance of means to ends (iv) seems even more rare: there is little evidence that non-human animals can do this (Shettleworth, 1998; but see Akins and Zentall, 1998).

## 9 Social Learning and Genetic Takeover

Now, how far are these different modes of transmission suited for the role I have ascribed to 'social learning'—that is, spreading complex adaptive behaviours from individuals to populations, and thereby rendering those behaviours available for 'genetic takeover'? There are immediate problems with all but the last. Stimulus enhancement (i) and goal emulation (ii) seem ill-suited for transmitting complex behaviours, while there is nothing in blind mimicry (iii) itself to favour the transmission of adaptive over non-adaptive behaviour.

The trouble with stimulus enhancement (i) and goal emulation (ii), from our perspective, is that they don't transmit complex behaviours as such; rather, they transmit the environmental opportunity, so to speak, with the learner then using its own devices to exploit the opportunity. To see the problem, imagine that some unusual or lucky individual lights on some complex tool-using strategy with which to extract grubs from holes. Stimulus enhancement means that other individuals will be more likely to find themselves in the conditions where this behaviour would be rewarded; but this won't get these individuals performing the behaviour, if its complexity makes it unlikely that they will then randomly generate it. Again, goal emulation means that those observing the expert will learn that there are grubs in holes; but this won't get them performing any complex tool-using behaviour either, if nothing analogous is already present in their behavioural repertoire.

Blind mimicry (iii) suffers from a different problem. Here it is specifically the behaviour that is being transmitted, rather than the opportunity, and so a learner may well pick up some complex sequence of behaviours from a demonstrator. But there is nothing in blind mimicry to ensure that learners will preferentially copy good tricks rather than bad ones. To the extent that the behaviour is being picked up without any appreciation of what results it brings, it is as likely that useless patterns of behaviour will spread as useful ones. Blind mimicry on its own thus fails to provide a mechanism by which a good trick will spread throughout a population once acquired by one individual.

These difficulties with the first three modes of social learning are not insuperable. Perhaps the aimlessness of blind mimicry will be moderated if learners only persist with the copied behaviour if they subsequently find it psychologically rewarding. This will have the effect of keeping good tricks in the population—and making them available for further mimicry—and weeding out bad tricks. (Alternatively, learners may selectively mimic dominant or prestigious individuals—this too will discriminate in favour of advantageous cognitive strategies, to the extent that dominance and prestige depend on such strategies. Cf. Richerson and Boyd, 2004.)

Conversely, elements of blind mimicry might help overcome the limitations of the first two modes of social learning. Animals who are introduced to new opportunities by stimulus enhancement and goal



emulation will be more likely to find some complex way of exploiting them if they are disposed blindly to mimic elements of the behaviour of others who have adopted some such means.

In any case, it is not as if there is some absolute level of reliable social transmission which needs to be reached. There will be cases and cases. We are interested in the possibility of genetic takeovers of complex adaptive learned behaviours. Such genetic takeovers require that the behaviour be reliably transmitted. There will be contexts where the requisite threshold of reliability is ensured by some mix of the three kinds of social transmission discussed so far, even if they are less effective at doing this than might initially have been supposed.

Even so, it should be clear that genetic takeover of complex behaviour is far more likely among individuals that are capable of the final mode of social learning, that is, learning about means to ends. Here there will no problem of bad tricks being as likely to be copied as good tricks—individuals will pick up specifically those behaviours that they can see give rise to attractive results, not just any behaviours they observe, as with blind mimicry. Nor is there any barrier to the copying of complex behaviours—individuals will here adopt the specific strategies they observe in their behavioural models, and will not be left to their own devices to develop ways of exploiting copied opportunities, as with stimulus enhancement and goal emulation.<sup>9[9]</sup>

This suggests that, while there may be a relatively limited range of cases in other animals where complex behaviours come under a genetic control as a result of first being learned, there will have been ample opportunities for such ‘Baldwinization’ in our own recent hominid ancestry. Perhaps I am being unduly negative about other animals here: the points raised in this section by no means fully rule out the possibility that genetic takeover has often played a significant role in cognitive evolution outside the recent hominid lineage. But, be that as it may, our main topic in this paper is human cognition, and the availability of explicit learning about means to ends among our recent ancestors means that they would not have faced the same barriers to the cultural transmission of complex behaviours as other animals.

## 10 Maladaptive Cultures

This emphasis on the explicit learning of means to ends, however, raises a rather different query about the genetic takeover of cultural practices. A cultural practice will be a candidate for genetic takeover just in case it is biologically advantageous. Genes that help you to learn P will be subject to natural selection just in case P increases reproductive fitness. By and large, we can expect learned behaviour to be so biologically advantageous—after all, learning mechanisms have been designed by natural selection to select reproductively advantageous behaviours in the light of experience. Still, such learning devices are not sure-fire, and in some environments they will end up selecting biologically non-adaptive behaviour.

This will be a particular danger with social learning via the explicit appreciation of means to ends. This is a highly sophisticated form of learning, which depends on the vagaries of individual experience in complex ways, and which therefore leaves plenty of room for biologically deleterious results. We need only think of the way that contemporary individuals socially acquire such habits as drinking alcohol, smoking, and piercing body parts. While there is a certain sense in which such behaviours are indeed ‘good tricks’—they are often genuinely effective means to feelings of well-being or to higher status—the social learning mechanisms of many individuals place far too much weight on these outcomes, and so instil behaviours which overall have a highly negative effect on reproductive fitness. And in such cases there will clearly be no question of genetic natural selection favouring genes which make you better at learning such behaviours, for the obvious reason that such genes will only decrease reproductive fitness even further.

Earlier in this section I argued that the explicit learning of means to ends is the mode of social learning most likely to facilitate genetic takeover. However, if this kind of social learning systematically gives

---

<sup>9[9]</sup> The reliable transmission of complex cognitive practices matters, not just for the possibility of genetic takeover, but also for the possibility of cumulative culture. This latter issue is the focus of Tomasello (2000). While Tomasello himself does not deny that the explicit appreciation of means-end relations matters for reliable transmission, he regards this as pretty much the same thing as the understanding of mind (and in particular, the identification of intentions). However, I think that non-human animals are blocked from an explicit appreciation of means to ends by far more fundamental cognitive barriers than their lack of understanding of mind. For discussion of this issue, see Papineau (2004) esp. sect. 7.

rise to biologically maladaptive practices, in the way just described, then this suggests that genetic takeover may not be a significant evolutionary process after all.

## 11 The Adaptivity of Vertical Cultures

The danger of biologically maladaptive cultural practices depends crucially on who learns from whom. In this connection it will be helpful to distinguish between ‘horizontal’ and ‘vertical’ transmission of cognitive practices. While horizontal transmission is indeed prone to pass on biologically maladaptive practices, this is not true of vertical transmission.

Horizontal transmission is perhaps the most familiar way of thinking about the promulgation of culture. Here individuals learn cognitive traits from other unrelated individuals—traits are passed ‘sideways’ from one individual to another, so to speak. When cultural transmission proceeds in this manner, cognitive traits will become prevalent the more efficient they are at so ‘infecting’ new individuals. Given this, such horizontal transmission does indeed open the way for biologically disadvantageous traits to spread.

However, an alternative mode of transmission is ‘vertical’, from parents to children. And here things work rather differently. To the extent that transmission is vertical, cultural traits will spread just in case they increase the reproductive success of their possessors. This is because vertically transmitted traits will thus be subject to a process of natural selection entirely akin to the selection of genes which contribute to individual reproductive success. So when transmission is vertical, only biologically advantageous traits will spread through a population. Vertical transmission is thus likely to create conditions that will foster genetic takeover after all.

It is somewhat unusual to think of cultural traits as subject to the same selective pressures as genes. There is plenty of literature, of course, which treats cultural traits as ‘replicators’ in their own right, as ‘memes’, in Richard Dawkins’ terminology (Dawkins, 1976, Blackmore, 2000). But most ‘meme’ theory focuses on horizontal transmission, and therefore views memes as being subject to quite different selective pressures from genes. With vertical transmission there is no such contrast, however. To the extent that cultural traits are passed from parents to children, they will be inherited in just the same manners as genes, and so are subject to entirely analogous selection processes.

Doesn’t the idea that ‘cultural traits are inherited in just the same manners as genes’, as I just put it, run counter to a central plank in modern biological thinking, namely, that only genotypes and not phenotypes are passed down from parents to children. Surely this is the central message of Waismann’s famous diagram: parental genotypes influence children’s genotypes, but parental phenotypes per se have no effects on children. A proficient hunter may become expert at throwing spears, but this doesn’t mean his children will automatically inherit this efficiency. However, it is easy to be misled by Waismann’s diagram. It is of course true that parental phenotypes do not influence children’s phenotypes by altering children’s genotypes. There is no downwards causation from phenotype to germ line (cf. Crick’s ‘central dogma of molecular biology’). But it does not at all follow that parental phenotypes do not influence children’s phenotypes at all. For there remains the possibility that they influence them directly, rather than by altering the germ line. And once this is in clear focus, then it is surely uncontroversial that phenotypes can indeed so be passed down from parents to children. The expert hunter’s proficiency will make no difference to his children’s genotypes. But it may make plenty of difference to their phenotypes, if they learn their hunting techniques from him.

Biological evolution by natural selection requires heritable traits. But there is no obvious reason why it should require this inheritance to be genetic rather than non-genetic. It is arguable that the promulgation of non-genetically inherited traits via their differential influence on reproductive success is just as much biological evolution by biological natural selection as more familiar cases of genetic evolution.

It is common, even among those who regard themselves as opposed to a gene-centred view of evolution, to allow that a change in gene frequencies is a necessary and sufficient condition for biological evolution, if only at a ‘book-keeping’ level. I am suggesting that even this is too much of a concession to gene-centrism. To digress for a moment, consider Matteo Mameli’s fable of ‘the lucky butterfly’ (2004). Suppose that there is a species of butterfly that imprints on the plant it hatches on. Butterfly larvae retain some trace of these plants, and when it is time for the mature butterflies to lay

eggs, they return to the plants they hatched on. The tendency to lay eggs on a given type of plant is thus non-genetically inherited, passed from mothers to offspring via this imprinting mechanism. Now suppose that some population of these butterflies are all imprinted on plant type A. Then a freak accident—a storm, say—leads one butterfly to deposit her eggs on a plant of type B. Because of the imprinting mechanism, her descendants henceforth lay their eggs on plant B. Suppose plant B is more nutritious, with the consequence that the descendants of ‘the lucky butterfly’ start outcompeting the other butterflies in the general Malthusian struggle for survival. After a while the population consequently comes to consist entirely of these descendants. I say, following Mameli, that this is a standard case of biological evolution by natural selection, even though there need have been no change whatsoever in the butterfly population’s gene pool. This might seem strange, but compare the scenario just outlined with one where the plant preference is indeed genetic, and the ‘lucky’ butterfly undergoes a genetic transformation that switches her from plant A to plant B, with the result, as before, that her descendants come to exhaust the population. There is of course no dispute that this would be biological evolution—one allele is favoured over another because of its advantageous effects. But if this would be biological evolution, why not regard the original lucky butterfly scenario in the same light? Why is it significant that the plant preference is determined by a gene rather than a memory trace, given that both are equally passed on from parents to offspring?

The story of the lucky butterfly is an artificial example. But there is reason to suppose there is plenty of reliable non-genetic inheritance in nature and that in consequence there is natural selection of the non-genetic traits so inherited. Mameli (2004) lists many real-life examples. In addition to imprinting for locality, as in the lucky butterfly fable, he considers various other kinds of imprinting, including imprinting for kind of habitat and imprinting for food and sexual preferences. More generally, he points out that less channelled forms of learning than imprinting also lead to offspring matching their parents in various respects. In the non-psychological realm, too, there are plenty of examples: various non-genetic zygotic materials are acquired by offspring from their parents, as are many symbionts. (See also Jablonka and Lamb, 1995, Avital and Jablonka, 2001.)

Mameli does not consider human cognitive traits. This is because the selective processes operating on these are complicated by the possibility of horizontal as well as vertical transmission. However, there is evidence that the vertical transmission of cognitive traits is an important mechanism among humans in traditional societies (Hewlett and Cavalli-Sforza, 1986, Gugliemino et al., 1995.) Perhaps this is itself the product of selection pressures operating within hominid history. Given that vertically transmitted traits will become common just in case they are reproductively advantageous, there will be extra genetic selection pressure in favour of learning from parents as soon as there is any tendency for selection to start operating on vertical transmission channels—for, once there is any such tendency, then genes which lead offspring reliably to copy their parents rather than other individuals will be favoured precisely because they are more likely to engender reproductively advantageous practices. (Cf. Laland et al., 2000, 142.)

I trust it is now clear how the selection pressures acting on vertically transmitted cognitive traits will favour reproductively advantageous traits, where this is not necessarily so for horizontally transmitted traits. Perhaps the point is most easily seen by considering the analogous pressures on parasites and symbionts. ‘Infectious’ parasites that are good at ‘jumping sideways’ may well be malignant to their hosts, for their long-term success is compatible with the reduced fitness or even death of those temporary hosts. But symbionts who spend extended periods of time in a single host, and whose descendants live in the offspring of that host, will outcompete their conspecifics just in case they help their hosts to survive and reproduce, for this is necessary condition for their reproductive success. Just the same applies to cognitive traits. Practices that spread sideways can be selected even if deleterious—like smoking, drinking and body-piercing. But practices that are transmitted from parents to children will spread only if they increase reproductive fitness, for their fates are bound up with the fate of the host lineages they inhabit.

Let us return from this digression into the biology of vertical transmission to our main topic, namely, the genetic takeover of complex cognitive practices. In section 9 I observed that the social learning of complex cognitive practices calls for explicit learning of means to ends, as opposed to less sophisticated forms of social learning. This then led in section 10 to the difficulty that there is no guarantee that this form of social learning will promulgate reproductively advantageous traits, as opposed to psychologically attractive ones. Explicit learning of means to ends is as capable of spreading unhealthy fashions as it is at instilling reproductive advantages. And if cognitive practices

are reproductively unhealthy, then there will be no question of genetic takeover—genes get selected if they help foster traits that yield reproductive success, not traits that are psychologically attractive.

However, we are now in a position to see that this worry about the possibility of genetic takeover disappears if the primary mode of transmission of cognitive traits is vertical. True, explicit learning of means to ends will still be capable of leading offspring to copy psychologically attractive but reproductively disadvantageous practices from their parents. But if these traits are transmitted vertically, then any such reproductively disadvantageous traits will tend to disappear, through failure of their possessors to have descendants onto which to pass them. Vertical transmission ensures that only reproductively advantageous traits will become prevalent throughout the hominid population. And therewith vertical transmission will create the conditions for genetic takeover. Since vertical transmission ensures that prevalent practices will be reproductively advantageous, it also means that genes that foster those practices will have a selective advantage.

## 12 Environmental Stability

Back in Section 5 we noted a different kind of requirement for genetic takeover. If there is to be an advantage to genes that bring some cognitive practice under genetic control, then the environmental conditions that make those cognitive practices advantageous must remain stable over evolutionarily significant periods of time. Since genetic takeover reduces plasticity, there will be no selection for genetic control if the relevant environments are variable.<sup>10[10]</sup>

It might seem that this argues against genetic takeovers of cultural practices in our recent hominid ancestry. *Homo erectus* and *Homo sapiens* are among the most adaptable species that have ever existed, managing to establish themselves in a very wide range of environments offering many different kinds of exploitable resources. Techniques of hunting, foraging, and defence that work in one such environment will tend not to work in others. To the extent that hominid lineages experienced variable environments, we would thus not expect such techniques to come under genetic control. Our ancestors would surely have done much better to tailor these techniques to current environments in the light of experience, rather than committing themselves genetically to particular strategies. (Cf. Sterelny, 2003, ch. 9.)

However, these considerations do not necessarily apply to all adaptive hominid cognitive practices. If we focus on specific techniques for food-gathering and defence, involving specific weapons, tools and techniques, then the variability of the relevant natural environments may well have prevented any genetic takeovers. But the same point does not apply to such general cognitive powers as linguistic capacity, understanding of mind, folk physics, folk biology, and so on. The advantages of these cognitive powers will not be tied to some specific environmental condition, but will rather be available across all natural environments. These general cognitive powers enhance access to information, increase understanding, and facilitate social coordination, and this will be of benefit to the possessors of such powers in any human society, whatever the natural environment.

In effect, the prior existence of a culture of these general cognitive practices is the only environment required to give a selective advantage to genes that will accelerate the learning of that culture. If a cognitive practice is advantageous in all natural environments, as language capacity, understanding of mind, folk biology and folk physics arguably are, then a culture that renders that practice learnable will itself constitute the environment that favours genes that lighten the learning load involved. As long as that culture persists, then those genes will be advantageous, increasing the reliability of learning and reducing the costs involved. Here genetic takeover does not depend on any stability of external, natural environments—it is enough that there is a cultural environment to provide a stable backdrop for the selection of the relevant genes.

---

<sup>10[10]</sup> Now that we have distinguished social from individual learning, we can add a wrinkle: social learning, rather than individual learning, will be advantageous when environments have an intermediate degree of stability, between the long-term stability that favours genetic control and the very low level of stability that demands individual learning. Social learning is less costly than individual learning, so will be better than individual learning when environments aren't so variable as to require recalibration of traits to circumstances in each individual's lifetime; at the same time, it is more flexible than genetic control, and so will be favoured when environments don't display long-term stability over multiple generations. (Boyd and Richerson, 1985, Laland et al., 2000.) From the point of view of the argument of this paper, however, we can lump all learning together as suited to 'variable' environments, given that our focus is on scenarios with the very high degree of environmental stability required to favour genetic control over even social learning.

### 13 Reliable Transmission

How stable are human cultures? Dan Sperber (1996) has stressed the point that cultural transmission is markedly less reliable than genetic transmission. Where sexual reproduction standardly transmits perfect copies of parental alleles, mutations aside, cultural transmission is subject to all manner of bias and noise. This argues that human cultures are unlikely to remain stable over significant periods of biological time, thus undercutting the last section's suggestion that such cultures can themselves provide the stable environments which will allow the selection of genes

Here it is worth distinguishing between vertical and horizontal transmission once more. In the section before last I observed that vertical transmission seems to play an important role among humans. Of course, there will be cases and cases. Humans certainly learn many things from individuals other than their parents. We need only think of contemporary adolescent teenagers, who generally regard their parents as absolutely the last people to adopt as role models. But this is consistent with the possibility that maturing humans acquire large amounts of cultural information specifically from their parents at earlier stages of their development. In particular, this seems likely for their acquisition of such basic cognitive powers as language capacity, understanding of mind, and so on. To the extent that these powers depend on cultural training, the most likely context is surely interaction between parents and maturing offspring.

If this is right, then it gives reason to suppose that the relevant cultures practices will constitute stable traditions rather than transient fashions. I have already observed that the prevalence of vertical transmission can create selective pressures in favour of genes that lead offspring to copy their parents. Let us now add in the point that, in cases where transmission is vertical, there will also be pressures for genes that lead parents to teach their children. This is a kind of mirror image of 'Baldwinization'. Here the possibility of offspring acquiring some practice via learning leads to the selection of genes which makes this learning more reliable and less costly—but here the genes operate through the parental teachers rather than the maturing learners.

Note that this latter possibility is peculiar to vertically transmitted learning. There is nothing in principle to rule out the genetic takeover of horizontally transmitted cultures. True, as we have seen, there is a question of whether horizontally transmitted practices will be biologically advantageous to their practitioners. And there is the issue, which we are presently discussing, of whether horizontal culture will be stable enough to sustain genetic takeovers. Still, as I say, there is no principled barrier to some horizontal culture satisfying both these requirements, and so yielding selective pressures for genes which make tyros better at learning the relevant practice. Yet even in such a case there will be no pressure for genes for teaching the practice, given that the transmission is horizontal. Since the beneficiaries of teaching will be unrelated individuals, it will not increase the teacher's reproductive fitness that the learners should acquire the trait. It works differently with vertical transmission. In that case, the beneficiaries of teaching will be the offspring of the teacher, and so parental genes which make the offspring better at acquiring the advantageous practice will automatically be favoured by natural selection.

A familiar example of this kind of selection is the way in which parents in many species will help offspring to practice their food-gathering skills (cf. Avital and Jablonka, 2001, 307-9); for example, some species of mammals and birds offer captive prey for their offspring to practice with. Clearly here there has been genetic selection on the parents for behaviour that facilitates learning in their offspring. It seems highly plausible that the natural tendency of human parents to engage in sustained verbal and intellectual interaction with their children, even at an age when the children cannot respond in kind, is similarly the product of selection of genes to make their children learn better.

Taken together, these points give us reason to suppose that the vertical transmission of basic cognitive practices like language, understanding of mind, and other folk theories would have been highly stable. Offspring would be naturally predisposed to copy their parents, and parents would be naturally disposed to teach their children various specific practices. These factors would seem quite adequate to sustain cultures in place for biologically significant periods of time, and thereby ensure the stable environments required for genetic takeover.

### 14 Innateness Revisited

Let us now finally return to the issue of the innateness of cognitive capacities. The argument of this paper may seem to support the thesis that many of our basic cognitive powers are innate. After all, it has aimed specifically to explain how the process of genetic takeover can lead to increased genetic control of practices that were previously learned. However, the issue of innateness is not straightforward. As I shall now show, in one good sense of ‘innateness’, there is no good reason to suppose that any cognitive capacities ever become innate.

How exactly is the notion of innateness best understood? One weak, comparative way of understanding the notion is in terms of ‘norms of reaction’. In this sense, a given genome makes some phenotype P innate to the extent that it ensures its appearance across a given range of environments. Accordingly, one phenotype will be more innate than another phenotype, relative to some genome, if it appears across a greater range of environments; similarly, one genome can make a given phenotype more innate than another genome would if it ensures the phenotype’s appearance over more environments. On this comparative understanding of innateness, there is no doubt that genetic takeover makes phenotypes ‘more innate’ than they were previously. By reducing the amount of learning required to produce some phenotype, genetic takeover means that the phenotype will appear in environments involving only limited amounts of learning, as well as environments involving more extensive learning.

However, this notion of innateness is limited to comparative judgements (and moreover will be rarely applicable, given that it requires the ranges of environments being compared to be related by strict inclusion rather than mere overlap). Because of this, many theorists aim for a stronger notion. One attractive notion is that a phenotype is innate just in case its appearance in normal development does not depend on any psychological mechanisms, and in particular does not depend on any learning process. (Cf. Samuels, 1998, 2002, 2004. See also Cowie, 1999.) This proposal is not unproblematic, facing obvious difficulties in its right-to-left direction: it is by no means clear that appearing in normal development without the help of any psychological mechanisms is sufficient for innateness. (Cf. Mamerli and Papineau, forthcoming, sect. 4.) However, we can by-pass this issue here, since I shall only be concerned with the converse left-to-right claim: something is not innate if it is produced by a psychological mechanism like learning. This seems relatively uncontroversial, and will plausibly part of any non-comparative notion of innateness. What I now want to argue is that there is no reason to suppose that genetic takeover will ever lead to innateness in any such non-comparative sense, on the grounds that it is unlikely ever to replace learning entirely by genetic control.

By way of an illustration of this point, consider the Galapagos woodpecker finches once more. Here there is no question but that their tool-using behaviour is in a comparative sense highly innate. Very little in the way of environmental support is needed for the behaviour to emerge. In particular, the finches seem not to need demonstrations by existing adepts from which to copy the behaviour socially. Even so, genetic control has not entirely eliminated the need for learning. The birds still need to be able to experiment with twigs at a crucial stage in development, in order to move from a crude predisposition to fiddle with twigs to successful insect-catching. It takes a month or two for the juvenile birds to refine this skill via individual trial-and-error learning (Tebbich et al. 2001). Their genes may strongly predispose them to the behaviour, but its full emergence also hinges on learning-based informational input from the environment.

A similar phenomenon is displayed in Hinton and Nowlan’s (1987) simulation. As I explained earlier, their simulation showed that, once their neural nets could learn to set their connections at ‘1’ rather than ‘0’, then the overall advantageous phenotype of all 20 ‘1’s became accessible, and genes for ‘1’ started being selected for, replacing the alternative alleles that left the connections to learning. However, Hinton and Nowlan’s simulation did not lead to the total replacement of learning genes by those that fixed ‘1’s without learning. Once the neural nets had something like 70% of their connections fixed by genes (with the exact percentage depending on the parameters of the specific simulation), then the selective pressures tailed off, and there ceased to be any significant further replacement of learning alleles. This was because it was a relatively easy task to learn to set the last few connections at ‘1’, once most of the others were genetically fixed at ‘1’, so at that stage extra genetic control ceased to be significantly advantageous.

There is a principled reason why genetic takeovers should display this kind of incompleteness, always leaving some role for residual learning. In order for genetic takeover to be possible at all, it cannot be too hard to learn the overall advantageous phenotype at the early stages when very little is genetically

fixed. If there were no real chance of finding the phenotype via learning in these early stages, then genes that marginally lightened the learning load would not be favoured, for they would still leave the organism with little chance of finding the pay-off phenotype. (This, recall, was why I attached so much significance to social learning. The point of social learning was that it can make complex behaviours learnable even when they are beyond the reach of individual trial-and-error learning.)

So candidate phenotypes for genetic takeover cannot be too hard to learn, even when they have little genetic help. An obvious corollary is that they will become very easy to learn, once there is a significant amount of genetic help. At that stage there will be no marked advantage to continued genetic takeover. Why bother to write the last details into the genes, when they can be picked up with no significant effort from the environment? Moreover, there may well be loss-of-flexibility costs associated with further genetic control, in the form of inability to fine-tune the phenotype to detailed environmental contingencies. All in all, then, it seems only to be expected that genetic takeovers will characteristically remain incomplete, always leaving some role to learning in fixing the overall phenotype. And to the extent that 'innateness' implies an absence of learning, this will mean that those phenotypes are never innate.

### 15 Learning all the Way Down

An obvious retort to this line of argument is that it may show that advantageous overall phenotypes are never rendered fully innate, but that this does not mean that components of those phenotypes will not be fully innate. Thus consider my baby model from section 4: I decomposed some overall phenotype P into components  $P_i$  each of which could be fixed by some allele  $G_i$  or alternatively left to learning by  $L_i$ . The argument of the last section gives us reason to doubt that the overall phenotype P will ever become fully innate, since the selective pressure to bring the last  $P_i$ s under genetic control will tail off. But this does not mean that none of the component  $P_i$ s will be fully innate—and indeed my model assumes that they will be, whenever the specific  $G_i$ s that fix them are present.

More generally, this is the natural line of response for anybody concerned to defend a strong cognitive nativism. Nobody, I take it, wants to argue that learning is unnecessary for the acquisition of natural languages, like English or Swahili, or for knowledge of specific biological categories, or even for the culturally variable elements of folk psychology and folk physics. Rather, according to nativist orthodoxy, it is the structures that facilitate these mature accomplishments that are innate, not the mature accomplishments themselves. Of course the full flowering of these accomplishments depends on some degree of learning. But this learning is made possible by some underlying structure (by some specialized learning mechanism, so to speak) which is itself fixed by the genes, and which owes nothing to informational input from the ontogenetic environment.

Within the classical computationalist tradition, this view gets cashed out as the claim that there are various bodies of innate knowledge. Since individuals get these bodies of knowledge from their genes, they do not need to extract them from their environments. Given this headstart, they are then in a position to learn the further items of information needed to complete the relevant capacities. Thus, 'universal grammar' is the innate body of knowledge that allows the acquisition of natural languages; similar innate bodies of universal knowledge are posited to account for the acquisition of folk psychology, biology and physics. Nor need this model be restricted to the classical computationalists. Connectionists will talk about prewired connection strengths, rather than innate sentences in the language of thought. But in the present context of argument this is not a substantial difference. There is nothing to stop us viewing connectionist prewirings as themselves embodying items of information, indeed just the same items of universal knowledge as are posited by classical computationalists.

I am happy to agree that, at some level of description, the genes that have been selected to foster specific cognitive capacities can be viewed in this way as fixing various elements of 'innate knowledge'. After all, these genes will have been selected because they combine with inputs from learning to produce mature cognitive phenotypes; given this, the 'informational content' of the genes can be equated with the inference from the learning input to the informational contents of those mature phenotypes. What remains open, however, is whether this kind of description will amount to anything recognizable as a component of linguistic knowledge, folk psychology, or other familiar cognitive accomplishment. For it may be that the contribution of the genes takes place at a very basic developmental level, altering neonatal perceptual saliences and building certain kinds of fundamental neural structures, with the construction of mature cognitive capacities requiring informational learning

input at every stage from then points onwards. If this is right, then learning as well as genes will be implicated in the acquisition of even the components of mature cognitive capacities, like the folk psychological ability to judge who can see what, or the linguistic disposition to identify anaphoric constructions, or the folk biological assumption that organisms have species-typical essences, and so on.

My earlier model of genetic takeover involving  $P_i$ s and  $G_i$ s was too restrictive in this respect. There I assumed that the overall phenotype  $P$  could be divided into recognisable phenotypic components  $P_i$ , each of which could either be entirely fixed by genes  $G_i$  or could be left to learning. But there was no essential reason, apart from expository simplification, to think of genetic takeovers in this way. Genetic takeovers require only that there are  $G_i$ s which lighten the learning load somehow, not that they do this by each fully determining some perspicuous component of the phenotype. The process would work just as well even if each such salient component were a product of both genes and learning, provided the genes involved did something to make it easier to learn the overall phenotype. (Cf. Papineau, 2006, section 6.)

The point generalizes to real-life examples. To see this, note that the considerations rehearsed in the previous section will apply as much to the salient components of any cognitive capacities as to the overall capacities themselves. Consider, as above, the folk psychological ability to judge who can see what, or the linguistic disposition to identify anaphoric constructions, or the folk biological assumption that organisms have species-typical essences, or so on. On the assumption that these abilities are upshots of genetic takeover, then they were once derived from ancestral learning mechanisms, and only subsequently has there been selection of genes to foster them. Given this scenario, there seems no reason to suppose that the genes so selected would have entirely eliminated any role for learning in the production of even these components. As before, given that ancestral learning was feasible, and the environment required available, why would selection have bothered, so to speak, to render these components fully innate? The selection of genes that make such learning fast and easy is one thing; the selection of genes that replace learning altogether is another.

To urge this is not to deny that genes resulting from genetic takeover will have some effects independently of contributions from learning. Moreover, as explained above, I have no objection to characterising these innate effects in informational terms, as items of 'innate knowledge'. However, to repeat the earlier point, there is no reason to suppose that these items of information will amount to anything recognizable as components of folk thinking. The fully innate effects of genes need not extend beyond the very earliest stages of development, fixing initial neural structures that bias learning in certain ways, but which from then on need to be combined with inputs from learning if further intellectual development is to occur.<sup>11[11]</sup>

It is a familiar general point that genes determine scarcely anything on their own, without some help from environmental factors: genes are selected to produce advantageous phenotypes in conjunction with stably recurring features of the environment. With those specific genes that result from genetic takeovers of previously socially learned practices, the relevant stable features of the environment will be the continued existence of that practice, which will then contribute via learning to the acquisition of that practice by maturing individuals.

The process of genetic takeover thus yields cognitive capacities which derive from a deep interaction between genes and learning. The striking ease and rapidity with which children master their native language and acquire various elements of folk thinking, even in the absence of any explicit instruction, provides undeniable evidence that many genes have been selected specifically to foster these cognitive capacities. However, to the extent that this selection has derived from genetic takeovers of ancestral cultural practices, then no recognizable component of these capacities is likely to be innate, in the sense that it would appear even in the absence of any learning. Genetic natural selection will have ensured that such capacities emerge quickly and reliably across a wide range of human environments. But since all human environments, freak cases aside, contain ample opportunities for social learning, continued dependence on some modicum of such learning will not detract from the speed and

---

<sup>11[11]</sup> Animal studies suggest that mature phenotypes often depend on earlier learning in unexpectedly deep ways. Young chicks who are prevented from seeing their own feet for two days after hatching are later unable to pick up mealworms, a typical behaviour in normal chickens (Wallman, 1979); mallard ducklings need to hear their own embryonic calls while still in the egg in order to recognize maternal mallard calls later (Gottlieb, 1997); rhesus monkeys reared in isolation are incapable of sexual behaviour when adult (Mason, 1960, 1961).



reliability of acquisition. In short, while genetic takeover selects genes for cognitive capacities, it does not make those capacities innate.<sup>12[12]</sup>

---

- Akins, C. and Zentall, T. 1998. 'Imitation in Japanese Quail: the Role of Reinforcement of Demonstrator Responding', Psychonomic Bulletin and Review 5, 694-7.
- Avital, E. and Jablonka, E. 2001. Animal Traditions: Behavioural Inheritance in Evolution. Cambridge: Cambridge University Press
- Baldwin, J. 1896. 'A New Factor in Evolution', The American Naturalist 30, 441-51, 536-53.
- Barkow, J., Cosmides, L., and Tooby, J. 1992. The Adapted Mind. New York: Oxford University Press.
- Bateson, P. 2004. 'The Active Role of Behaviour in Evolution', Biology and Philosophy 19, 283-298.
- Blackmore, S. 2000. The Meme Machine. Oxford: Oxford University Press.
- Boyd, R. and Richerson, P. 1985. Culture and the Evolutionary Process. Chicago: Chicago University Press.
- Chomsky, N. 1972. Language and Mind. New York: Harcourt, Brace, and World.
- Chomsky, N. 1988. Language and Problems of Knowledge. Cambridge, Mass: MIT Press.
- Cowie, F. 1999. What's Within? Nativism Reconsidered, New York: Oxford University Press.
- Dawkins, R. 1976. The Selfish Gene. Oxford: Oxford University Press.
- Dawkins, R. 1996. Climbing Mount Improbable. London: Penguin Books.
- Dennett, D. 1991. Consciousness Explained. London: Allen Lane.
- Fodor, J. 2000. The Mind Doesn't Work that Way. Cambridge, Mass: MIT Press.
- Godfrey-Smith, P. 2003. 'Between Baldwin Scepticism and Baldwin Boosterism', in Weber, B. and Depew, D. (eds), Evolution and Learning. Cambridge, Mass: MIT Press.
- Gottlieb, G. 1997. Synthesizing Nature-Nurture: Pre-Natal Roots of Instinctive Behaviour. Mahwah NJ: Lawrence Erlbaum.
- Griffiths, P. 2006. 'The Baldwin Effect and Genetic Assimilation', in P. Carruthers, S. Laurence and S. Stich (eds) The Innate Mind: Culture and Cognition. Oxford: Oxford University Press.
- Guglielmino, C., Viganotti, C., Hewlett, B. and L. Cavalli-Sforza. 1995. 'Cultural Variation in Africa: Role of Mechanisms of Transmission and Adaptation', Proceedings of the National Academy of Sciences 92: 7585-7589.
- Hewlett, B. and L. Cavalli-Sforza. 1986. 'Cultural Transmission among Aka Pygmies', American Anthropologist 88: 922-934.
- Hinton, G. and Nowlan, S. 1987. 'How Learning can Guide Evolution.' Complex Systems 1, 495-502.
- Jablonka, E. and Lamb, M. 1995. Epigenetic Inheritance and Evolution. Oxford: Oxford University Press.

---

<sup>12[12]</sup> For helpful responses to earlier versions of this paper, I would like to thank Ruth Kempson, Richard Samuels, Gabriel Segal, and especially Matteo Mameli.

- Laland, K., Olding-Smee, J., and Feldman, M. 2000. 'Niche Construction, Biological Evolution, and Cultural Change', Behavioural and Brain Sciences 23, 131-75.
- Laurence, S. and Margolis, E. 2001. 'The Poverty of the Stimulus Argument', British Journal for the Philosophy of Science 52, 217-76.
- Mameli, M. 2004. 'Nongenetic Selection and Nongenetic Inheritance', British Journal for the Philosophy of Science 55, 35-71.
- Mameli, M. and Papineau, D. Forthcoming. 'The New Nativism: A Commentary on Gary Marcus's Birth of the Mind', Biology and Philosophy.
- Mason, W.A. 1960. 'The Effects of Social Restriction on the Behavior of Rhesus Monkeys: I', Journal of Comparative and Physiological Psychology 53, 82-9.
- Mason, W.A. 1961. 'The Effects of Social Restriction on the Behavior of Rhesus Monkeys: II', Journal of Comparative and Physiological Psychology 54, 287-290.
- Papineau, D. 2004. 'Human Minds', in A. O'Hear (ed.) Minds and Persons. Cambridge: Cambridge University Press.
- Papineau, D. 2005. 'Social Learning and the Baldwin Effect' in A. Zilhao (ed) Rationality and Evolution. London: Routledge.
- Papineau, D. 2006. 'The Baldwin Effect and Genetic Assimilation: Reply to Griffiths' in P. Carruthers, S. Laurence and S. Stich (eds) The Innate Mind: Culture and Cognition. Oxford: Oxford University Press.
- Pinker, S. and Bloom, P. 1990. 'Natural Language and Natural Selection', Behavioural and Brain Science 13, 707-84.
- Richerson, P. and Boyd, R. 2004. Not in Our Genes Alone. Chicago: University of Chicago Press.
- Samuels, R. 1998. 'What Brains Won't Tell Us about the Mind', Mind and Language 13, 548-70.
- Samuels, R. 2002. 'Nativism in Cognitive Science', Mind and Language 17, 233-65.
- Samuels, R. 2004. 'Innateness in Cognitive Science', Trends in Cognitive Science 8, 136-41
- Shettleworth, S. 1998. Cognition, Evolution and Behavior. Oxford: Oxford University Press.
- Sperber, D. 1996. Explaining Culture. Oxford: Blackwell.
- Sterelny, K. 2003. Thought in a Hostile World. Oxford: Blackwell.
- Tebbich, S., Taborsky, M., Fessl, B., and Blomqvist D. 2001. 'Do Woodpecker Finches Acquire Tool Use by Social Learning?' Proceedings of the Royal Society 268: 2189-2193.
- Tomasello, M. 2000. The Cultural Origins of Human Cognition. Cambridge, Mass: Harvard University Press.
- Turney, P., Whitely, D., and Anderson, R. (eds). 1996. Evolutionary Computation, Evolution, Learning and Instinct: 100 Years of the Baldwin Effect. Cambridge, Mass: MIT Press.
- Waddington, C. 1953. 'Genetic Assimilation of an Acquired Character', Evolution 4, 118-26.
- Waddington, C. 1957. The Strategy of the Genes. London: Allen and Unwin.
- Waddington, C. 1961. 'Genetic Assimilation', Advances in Genetics 10, 257-90.

Wallman, J. 1979. 'Animal Visual Restriction Experiment: Preventing Chicks from Seeing Their Feet Affect Later Responses to Mealworms', Developmental Psychobiology 12, 391-7.

Watkins, J. 1999. 'A Note on Baldwin Effect', British Journal for the Philosophy of Science 50, 417-23.

---