Innateness and Genetic Information

Peter Godfrey-Smith

Philosophy Department
Harvard University

2007 [R]. Final version.

Forthcoming in *The Innate Mind: Foundations and the Future*, edited by Peter Carruthers, Stephen Laurence, and Stephen Stich. Oxford University Press.

Abstract: The idea that innateness can be understood in terms of genetic coding or genetic programming is discussed. I argue that biology does not provide any support for the view that the whole-organism features of interest to nativists in psychology and linguistics are genetically coded for. This provides some support for recent critical and deflationary treatments of the concept of innateness.

1. Introduction

The concept of information, along with its relatives, has always played an important role in discussions of innateness. That role has recently become even more prominent, because of a convergence or allignment between trends in psychology and biology. In psychology, an information-processing or representational view of the mind has continued to shake off rivals, and seems to become more and more entrenched. Given this framework, it will be natural to describe much innate structure in informational terms. Similarly, biology in recent years has seen a striking proliferation in the use of informational concepts in discussions of genes and gene action. The sense in some quarters is that a sharper, more rigorous information-theoretic language has arisen, which can complement or supplant the blunter instrument of ordinary causal description. As a consequence, the idea of the genome as an alternative *source* for the information realized in the structure of an agent's nervous system has apparently been moving from the status

of hopeful metaphor to something like a literal and theoretically rigorous mode of description.

In recent philosophy, in contrast, the concept of innateness has been subjected to a good deal of criticism and skepticism (Cowie 1999, Griffiths 2002, Mameli and Bateson 2006). Even defences of the concept are taking on a deflationary tone (Samuels 2002, Cowie forthcoming). But perhaps the philosophical treatments have not fully realized the promise of the informational framework so prominent in psychology and biology; maybe the distinction between innate and acquired traits can itself be understood in informational terms. The innate traits of an organism are the ones that are coded for, represented, informationally specified, or programmed for, by the organism's genetic endowment. The characteristics that are not coded for (etc.) are not innate, but acquired. That possibility is the topic of this paper. The next section discusses recent philosophical treatments of innateness, and the remainder of the paper looks at whether and how the "informational turn" in biology might transform that discussion.

My conclusions are largely negative, and lend some support to the tradition taking a skeptical or deflationary attitude towards innateness. Concepts of information and coding promise to both allign with and make genetic sense of the idea of innateness. But the allignment, on close inspection, is poor, and we also have reason to treat much of what these concepts offer with suspicion.

2. Options for innateness

The *apparent* role of the concept of innateness, and the natural starting point for philosophical analysis, is that "innate" is used in higher-level sciences like psychology to pick out a natural class of phenomena that can be described in more detail by developmental biology. For any given trait seen in an organism, it can be asked whether it was innate or not. Psychologists (and linguists, etc.) will use their own evidence to assess such hypotheses, but in principle a more detailed answer could be given by biology. So a philosopher looking at the relations between adjacent sciences will naturally try to find out exactly which property or category recognized by biologists is being gestured towards by appeals to innateness.

Some recent discussion has pursued a project of exactly this kind, but there is also a growing tradition of work that is more critical of the concept of innateness. We might distinguish some of the options as follows:

Reduction of innateness to a natural kind recognized by biology: This is the familiar approach sketched above. The attempted analyses of innateness in terms of genetic information and coding discussed below fall into this category. Another view of the same kind is André Ariew's analysis of innateness in terms of *canalization* (1996). A canalized trait is one that appears reliably in some organism in the face of substantial variation in environmental circumstances.

Options of this first kind may, of course, be mildly revisionary of familiar ways of using the concept. For example, they may replace the binary "innate/acquired" distinction with a gradient one. But they see talk of innateness as picking out a real biological phenomenon with reasonable accuracy.

Cluster concept, or family resemblance: Mameli and Bateson (2006) treat innateness as a cluster concept, distinguishing about a dozen "I-properties" that have some link to it. It is an open question how well the I-properties are correlated. If they are poorly correlated, then the concept becomes problematic because it will make bad inferences seem natural. If they are very well correlated, then this option shades into the first one above.

A related "family resemblance" option is defended in a new treatment by Cowie (forthcoming), which is less critical of innateness than her *What's Within?* (1999). In the new work, Cowie suggests that innateness is a family resemblance concept that provides a good temporary label for something we are trying to study. But for Cowie, we can expect future sciences to use a different categorization.

Disciplinary marker: Samuels (2002) argues that to be innate is to develop in an individual via a process that does not have a description in psychological terms. Samuels does not try to give a theory-driven account of which mechanisms have a psychological description (for example, an account in terms of information processing). For all Samuels

3

says, the distinction could be one that has grown up to mark a largely arbitrary disciplinary boundary.

If Samuels' account was supplemented with a list of which processes and mechanisms are the proper subject matter of psychology, in virtue of their empirical characteristics, then his analysis would be brought closer to the first category above. If, on the other other hand, his account was supplemented with a more sociological account of the psychology/biology boundary, the resulting treatment of innateness would become deflationary in character. And although Samuels does not present himself as a deflater (see, for example, his chapter in this volume), some of his remarks on the key disciplinary boundaries seem to me to push in what amounts to the more deflationary direction.

A somewhat similar role for innateness-talk was postulated in Cowie's pioneering *What's Within?* (1999). Cowie argued that the categorization of a trait as innate is often an expression of metatheoretic pessimism, a gesture towards the impossibility of giving a naturalistic explanation for the trait's appearence. However, Cowie's 1999 analysis also regarded some talk of innateness as more substantive: innate traits are acquired by a *special-purpose* mechanism, especially one that contrasts with traditional domain-general learning mechanisms.

ElimiNativism: Perhaps the most hostile recent treatment is Griffiths (2002). He sees the concept of innateness as a product of an essentialist "folk biological" mode of thinking (Atran 1990, Carey 1985). Innateness is linked to some real empirical distinctions – here he gives a less detailed list of the kind seen in Mameli and Bateson. But innateness is a concept that tends to pull its users beyond low-key operationalizable distinctions, towards the idea of innate traits as expressions of an "inner nature," conceived with normative loading. However strenuously a scientist might try to use empirically defensible definitions when applying the concept, the folkbiological habits of thought associated with the idea of innateness act "as a sink that draws new stipulative uses back towards the established [essentialist] use" (p. 82). So the concept should simply be avoided. (The pun "ElimiNativism" is due to Cowie (forthcoming).)

The idea that innate traits are those encoded or informationally specified by the genes falls squarely into the first family of options above. This idea is also more ambitious than the other approach mentioned above in this category – Ariew's analysis in terms of canalization. The idea of a developmental process that is insensitive to environmental circumstances is clearly acceptable in itself, and involves only a low-level piece of biological theorizing. Canalization can be used in more theoretically loaded senses, but that is not what Ariew has in mind. He has in mind a standard, modern, low-key usage. So the concept is not far at all from an informal concept of insensitivity to environmental circumstances that would have been available, in principle, to scientists and philosopher from much earlier times. So if Ariew is right, what we have here is a fairly simple sharpening-up of an old concept. Because some such notion of insensitivity to environment was conceptually available well before the 20th century, if people in the past did *not* guide their usage of "innate" with this sort of criterion, then any present-day usage that *does* use such a test is use of a new or modified concept. So arguments against Ariew are mostly arguments against his semantic claims about how the existing concept of innateness relates to phenomena of mere insensitivity (see Samuels 2002, for example).

A treatment of innateness in terms of genetic coding has a very different status. This idea could not have been conceived of before genetics, so it could not be what philosophers and scientists in a pre-genetic era "really meant" by the term. Instead, the model here is something like the relation between the term "water" and H_2O. The stuff that people in earlier times were talking about when they said "water" is in fact H_2O. Similarly, perhaps, the features that people were talking about in the past when they correctly labelled something as "innate" were in fact genetically encoded.

My critique in this paper is directly mainly at the very idea that there is a class of whole-organism characeristics, including cognitive ones, that are genetically encoded or informationally specified by the genes. So in principle, someone might say that the term "innate" is *aimed* at picking out the genetically encoded traits, while agreeing with me that this class is nearly empty, at least with respect to the traits that are of interest to psychologists. But that would be an unusual combination of views. My own response to

what I see as the failure of the treatment in terms of coding or information, and of similar views, is to see some support for the deflationary options listed above.

3. Genetic coding and genetic information

My main target in this paper is the idea of a division of an organism's characteristics into those that are genetically encoded (represented, programmed...) and those that are not. So my focus is on the kind of thinking exemplified here by Frank Sulloway:

> [N]o one has identified any genes that code for altruistic behavior. Such genes are nevertheless believed to exist because certain aspects of personality that underlie cooperative behavior – for example, empathy, sociability, and even altruism itself – are moderately heritable. (Sulloway 1998, p. 34)

Those who treat genetic coding in this way may or may not use that idea to make sense of innateness. Some probably find "innate" a term with dubious connotations, and talk simply of genetic coding in order to avoid misunderstanding. Others, however, are happy to make the connection explicitly.

> [I]f Narrow UG [Universal Grammar] is innate, it must be coded genetically, just like any specialized cognitive capacity in any animal, such as bat sonar. (Cullicover and Jackendorf 2005, p. 13)

In this section I will mostly discuss the idea of coding for traits itself, but will occasionally also comment on how this idea relates specifically to the concept of innateness.

I assume, at the outset, that when people talk of a characteristic as genetically coded (or represented, or programmed for...) they do not mean these phrases as mere synonyms for *genetically caused* or *genetically determined*. It is widely accepted that it is very hard to make sense of genetic causation in a way that will yield a suitable distinction between traits that are caused by (or mostly caused by) genes and traits that are not. Here

the arguments are familiar, and the discussion seems to have come to rest – for now – at a kind of low-key "interactionist" position. All traits are the causal product of both genetic and environmental factors. If we hold either side constant, we can make causal claims about the results of varying the other. And some traits are more sensitive than others to one or the other kind of variation, within a normal range or within a range realized within a given population at a specific time. But if the aim is a binary distinction between traits that do, and traits that don't, have genetic factors as "primary" causes, it is extremely hard to make sense of this, at least without drawing on population-wide facts that seem extraneous to the causal questions about a given individual's development (Lewontin 1974, Oyama 1985, Sober 1988, Block 1995, Cowie 1999). Part of the appeal of the idea of genetic coding is that it firstly seems to be a concept configured in a way that can yield a binary distinction, and secondly that it enables the proponent of innate traits to avoid accusations of crude genetic determinism.

So the idea of a genetically *coded* trait is not supposed to line up with the idea of a genetically *determined* trait. The genetic message, if there is one, need not always be reliably expressed. (We see this in the Sulloway quote above. A *moderate* level of heritability is taken to support the existence of a gene *coding* for altruism.) Neither is the idea of a genetically coded trait supposed to line up with the idea of a trait that is *affected by genetic causes* – that would be too weak, including just about everything about an organism.

I do not think that the place reached by the existing discussion of the potential role of ordinary causal concepts is conclusive. There may be more options here than have been supposed, using either ordinary causal concepts or carefully-honed relatives of them (Woodward 2003). But that is not the topic of this paper; the aim here is to assess the possibility of using informational or semantic concepts as an *alternative* to a direct appeal to ordinary causal relations, when understanding the innate as the genetic.

The next thing to do is to identify, and then put to the side, a concept of information that is undoubtedly useful in biology but which is also unsuitable for giving an analysis of innateness. This is the sense developed by Shannon (1948), and others, for use in the mathematical theory of information. In Shannon's sense, anything is a *source* of information if it has a number of alternative states that might be realized on a

particular occasion. And any other variable *carries information* about the source if its state is correlated with that of the source. This is a matter of degree; a signal carries more information about a source if its state is a better predictor of the source, less information if it is a worse predictor.

What Shannon did was provide a quantitative way of thinking about ordinary relations of physically-based correlation. This is very valuable in some contexts, but it does not – at least in its own – give us a sense of information which can do a better job than ordinary concepts of causation and correlation in making sense of innateness. Both genetic and environmental conditions can carry information about an organism's traits, if both of them covary with it. So this is not a notion of information that singles out a special *kind* of role that genetic factors have (Griffiths and Gray 1994). Information in Shannon's sense could be used, in some cases, to achieve a kind of apportioning of the roles of genetic and environmental causes. But what this would be telling us, in essence, is which kind of knowledge is *more informative for us*, when trying to predict what an organism will look like. Various other quantitative tools can be used to do this sort of apportioning. All of them, at bottom, are just measures of how genetic, environmental, and phenotypic variables in a given population are *associated* with each other.

So information in the Shannon sense is not something that has an in-principle association with genetic as opposed to environmental mechanisms, and the problem of apportioning of "informational influence" between genes and environment collapses (and messily) to the old problem of apportioning causal influence, which we are trying here to get beyond.

By now it should be a bit clearer what is being sought. The idea of coding (or representing, etc.) is supposed to make it possible for us to describe some special *kind* of natural connection between some genes and some traits, that can be used to underwrite innateness claims. The type of connection in question should at least approximate one that is either present or absent in a binary way. The connection in question will presumably have a basis in physical and causal facts of some kind, but this basis might be quite complex. (For example, it might be partly historical and partly present-tense.) It might be necessary to say that this special relation *supervenes* on lower-level physical facts, without being simply reducible to them.

The reader might suspect that I am setting up the project in a way that looks hopeless. In fact, though, I think that there is one set of relations between genes and one set of their causal products that *is* distinctive in a way that motivates a kind of semantic description (Godfrey-Smith 2000). This is found in the first few links in the causal chains between genes and their phenotypic effects. These are the causal processes by which genes template for amino acid sequences in the production of individual protein molecules. These are chemical processes with a normal physical basis, of course, but they have a highly distinctive character. The DNA sequence, via a largely fixed and combinatorially structured rule, acts as a template in the specification the linear order of amino acids in the formation of protein molecules. The DNA sequence is, in a sense, "read" by a complex of cellular machinery (including mRNA and tRNA molecules, and associated enzymes) in protein synthesis. The "genetic code," in its narrow and original sense, is the rule that maps DNA to amino acid triplets in this process.

So if we ask about the relations between a DNA sequence and a protein molecule, we find that it is, roughly speaking, a yes-or-no matter whether that DNA molecule codes for that protein or not. (The qualification "roughly speaking" has to do with the role of mRNA editing and other complications.) There is a (qualified) sense in which we can say, on the basis of biological theory, that the DNA present in a zygote contains a coded specification of all the proteins that the organism will make in its life. So there is nothing incoherent about the idea that *some* things about an organism are specified by a symbol-like structure that is present at birth; we see this (roughly speaking) in the case of the primary structure (linear sequence) of protein molecules.

So if "innate" means genetically coded, then an organism's lifetime stock of the kinds of protein molecules it will use in its cells is innate. (When I say this, I pay little heed to some biological details and qualifications that other philosophers make a great deal of – see, for example, Sarkar 1996 – that do complicate the situation. I think the basic picture survives the addition of those details.) However, this special form of description of gene action, in terms of products coded for, extends *only* to the formation of a protein molecule. After that, the protein goes off and joins the great causal soup of cellular processes, and has a role that is sensitive to the details of many other biological structures around it.

Is that all that can be said? The protein is coded for and then it joins the causal soup? Clearly more can be said. The causal role of the protein can be described in detail. We can say which reactions it catalyzes, what structures it helps to make, and how it might help regulate the production of other proteins by the genes. And the result of all this will be the highly reliable appearence of some whole-organism phenotypic traits, and the more sensitive and contingent appearence of others. Via such processes, cognitive machinery is built, along with immune systems and livers. But my point is that once we get past the production of each protein molecule, the role for a special "coding for" relation is over. The theoretically motivated form of semantic or symbolic description of genes is confined to low-level processes within individual cells. Roughly speaking, it is confined to the process of protein synthesis. (That claim will be qualified in the next section.) Once we are asking questions about levels of aggression, about universal grammar, about sexual orientation or basic numerical cognition, we have far outrun the kinds of entities that can be said, on the basis of biological theory, to be coded for. When understanding those sorts of traits, we are back in the land of ordinary causal explanation.

It may be objected, at this point, that surely there are some traits that are the relatively direct and straightforward consequences of the presence of specific protein molecules. The image of a "causal soup" is not applicable in all cases. Perhaps innate traits might include those that are the fairly direct products of a protein that has been coded for? If successful, an objection based on this line of thinking would enable the list of "genetically coded" characteristics to be expanded to include various genetic diseases, such as sickle-cell anemia, and other comparatively simple ways for the system to be broken (Moss 2003). We are still a very long way from the complex characteristics that nativists want to discuss.

At this point, it is possible for the debate to turn once more towards the problems with causal explanation that were put into the background earlier. Is there really nothing that can be made of the idea that some complex traits are reliably constructed via internal processes under genetic control, with only minimal and non-specific support from external factors? Perhaps there is something that can be made of it, but then we must grapple directly with the problem of attributing causal responsibility. My argument here

is that the notion of genetic coding, when it is well-motivated by the biology, does not itself contribute anything to the resolution of this problem.

The core of my position has now been laid out, but at this point it becomes necessary to look at a family of views in this area that might provide another way forward. A number of philosophers and biologists have argued that the ideas developed within the "teleosemantic" movement within philosophy of mind might provide the key to the semantic characterization of genes. When an entity has been subject to and shaped by a history of natural selection, this can provide the grounding for a kind of purposive or normative description of the causal capacities of that entity (Wright 1976, Buller 1999). That in turn can provide, according to teleosemantics, a basis for the ascription of semantic properties (Millikan 1984, Papineau and McDonald 2006). Various people have been attracted to the idea that genes, via their embedding in histories of natural selection, might have the function to represent or to carry information about their phenotypes (Sterelny, Smith, and Dickison 1996, Maynard Smith 2000, Shea forthcoming). An especially relevant feature of this approach is that if genes have such properties at all, the content of the genetic message would presumably involve whole-organism phenotypic characteristics, not mere protein molecules, because it is the whole-organism phenotypes that figure in histories of natural selection.

There are various ways in which the details of such an account can be developed, some focusing on the evolved functions of the genetic machinery as a *whole*, and others on the natural selection of *particular* genetic elements (Godfrey-Smith 1999). All versions of this idea offered so far have encountered problems. At the core of a number of specific problems is the fact that there is no simple overall connection between biological function, even in the rich historical sense, and having semantic properties. Legs are for walking, but they do not encode walking. Enzymes are for catalyzing reactions, but they do not represent this activity. There are things that legs and enzymes are *supposed* to do, but this does make them into representers or information-carriers. Why should it do so for genes?

Sterelny, Smith, and Dickison (1996) seem to think there is a quite intimate connection between evolutionary function and semantic properties in the case of biological structures that have been selected to play a causal role in *developmental*

processes. They argue that genes, in virtue of these functional properties, represent the outcomes they are supposed to produce. But once again, there are plenty of causal factors that have evolved to play a specific role in developmental processes without thereby acquiring representational content. The machinery *by which* gene expression occurs provides an example. This machinery is clearly involved in the same "developmental" processes that genes are.

If the mere invocation of biological functions, of the rich historical kind, does not solve the problem, then what was it that made teleosemantics so appealing to many people in the philosophy of mind? In the best-developed teleosemantic theory, that of Millikan (1984), the appeal to biological functions was made against a specific backdrop, which is relevant here. Millkan's theory has it that anything that has representational content functions as an intermediary between two devices, a "producer" and a "consumer." It is the role of the "consumer" that is most important here. I have argued elsewhere (2006) that we should think of the application of semantic and representational concepts in these sorts of scientific contexts in terms of the application of a causal schematism, or (in one sense) a model, that is derived from central cases of everyday symbol use. The most familar kind of symbol-using phenomenon in everyday life is one in which a person consults the state of one thing, X, to guide his or her behavior towards something else, Y. More simply, X is used as a "stand-in" for Y. So this is a picture in which something can only be a representation if there is a reader, consumer, or interpreter device of some sort. I think this picture or schema is at the heart of our habits of description of representations and symbols.

In naturalistic philosophical and psychological work on representation, people sometimes avoid this sort of picture, because of the obvious threat of regress and pseudo-explanation. Psychologists want to say that our minds contain inner representations. But does that mean they want to say that our minds contain little readers of the representations as well? If so, a regress threatens, as the symbol-using intelligence of the reader must now be explained as well. But the problem is best resolved without pretending that things can play the role of representations in the absence of anything like a reader at all. The problem can be resolved by noting that the familiar public phenomena contribute a causal schema that can be flexibly applied. Something can play a reader-like

role – well enough for the schema or model to be useful - without being anywhere near as intelligent as an agent. Bee dances represent (in a low-level sense) the location of nectar in part because of their relation to readers, even though the readers are very simple. The hardware of an ordinary computer also respects this distinction fairly well, in the separation between the processor and memory that is read or consulted. So the idea that representations involve readers need not set off philosophical alarm bells in quite the way that people often suppose. And Millikan's theory of meaning makes good use of this feature of our core ways of thinking about representations.

These general reflections are highly relevant to the case of genes. Let us think again about the cell-level processes of protein synthesis that were discussed earlier in this section. Here, I said, there is some real theoretical motivation for talking of a genetic code linking DNA sequences to amino acid sequences. We can now add some more detail about why that is so. In the transcription-translation processes that make up protein synthesis, there is a fairly clear sense in which a set of mechanisms that are physically distinct from the DNA message engage in a reading or consulting of it. There is not a single reader, admittedly, but a complex of elements (including ribosomes, RNAs, enzymes) that treat the DNA as a linear template in the construction of protein molecules. As with paradigmatic readers in everyday life, the rule of interpretation is relatively fixed, in comparison to the large variety of messages that are possible. If someone was to insist that the genetic message is not *literally* a message, not *literally* a code, because it is not properly embedded in a community of agents of the right kind, there is no real need to argue with this. One of the odd things about the description of semantic phenomena is that it is not at all clear where literal cases end and metaphorical cases begin. But the mechanisms underlying protein synthesis do, surprisingly enough, show a fairly close resemblance to paradigmatic cases of representation use, at least with respect to the "causal skeleton" of such phenomena.

With this point on board, we can note, from another angle, the crucial difference between the idea that genes code for proteins and genes code for whole-organism traits. Whereas there is a mechanism that has a kind of "reader" function in the creation of individual proteins, there is no organism-level reader of the whole genome, or anything even close to that. There *could* have been such a thing, in principle. We might have

evolved with hearts, eyes, livers, and a sort of "central reader" of the genetic code that coordinates our processes of development. But, of course, we did not. Each of our cells contains little readers, roughly speaking, but their job is done when individual proteins are produced. And crucially, a lot of little readers of this kind do not amount to a single big, organism-level reader, reading a grand message with a different and more complicated content.

One writer influenced by the teleosemantic tradition, Nicholas Shea, has attempted to confront these issues head-on (forthcoming). He accepts that representations require "consumers" of some kind, and accepts that the idea that genes contain a representation or specification of whole-organism traits requires a consumer over and above the cell-level readers that act in protein synthesis. His response is to argue that we do have good reason to recognize a kind of higher-level consumer of the genetic message, distributed across the organism. What Shea calls "inheritance systems" have the function to read the genetic message and produce heritable phenotypes.

Shea is right that *if* someone wants to use teleosemantic ideas to motivate a semantic interpretation of genes, and wants to claim that genes represent or encode whole-organism phenotypes, then something like this rather adventurous step is needed. Earlier work on the topic never grappled properly with the role of consumers or readers of the genetic message. But once we accept this fact, I think it becomes clear how tendentious the project of treating genes as encoding whole-organism phenotypes really is. In order to make such a view workable, Shea is forced to argue for the reality of something that is described like a piece of biological machinery, but which is in fact an abstraction. Shea sees his high-level consumer as a "system" bound together by a special kind of teleofunctional property (the function of producing "heritable gene-phenotype correlations to be acted on by natural selection"). Can an object like this interact with localized DNA molecules, or does it read a message that is itself another abstraction, something distinct from the molecules themselves? Do we have any independent motivation to believe in these things at all, other than their enabling us to say something extra and intuitive about the content of the genetic message?

4. Genetic programs

One last concept deserves a careful look here, and that is the idea of a *program* in the genes. This is less often discussed in philosophy (though see Moss 1992, Sterelny 2000), but it is a very common way of talking, at least informally, in the sciences. Here the link to the concept of innateness is not as potentially direct as it is in the case of coding. While a trait itself might be said to be coded for, talk of programs is most directly aimed at *processes*. But we could say something like this: characteristic C is innate in some organism if the genes in that organism program the developmental process to produce C.

In some ways the program concept does seem quite promising, in the light of the problems discussed in the previous section. It might be possible to argue that if we have a whole system made up of small parts that can each be described as executing a program, and the parts' activities are highly coordinated with each other, there *is* a sense in which the whole system is executing a higher-level program. So we might have a situation in which, at a low level of biological description, we have devices programmed to "make this amino acid chain here, now," and lots of these activities might be said to constitute the execution of a program at the level of the organism – a program for making a universal grammar facility, or an eye. The main problem with the notion of coding in this context was the need for there to be a suitable reader of any message that is said to exist. A whole-organism content needs a whole-organism message, which in turn requires an organism-level reader. The idea of a program seems different; a program existing at the level of the whole organism can be executed via myriad of low-level processes that act at a local level.

So the idea is certainly worth discussing. But now we need to look more closely at the key concepts themselves. When we take a coordinated set of causal processes of this kind and add the idea that they constitute the execution of a program, what are we saying? What are we adding to the idea that these are orderly causal processes that reliably produce some complex result of interest to us? In the biological case we cannot appeal to the role of the intentions of a human programmer. And even in the case of ordinary computers, of the kind on your desk, it is harder to answer this question than it might appear (Smith 2002).

I will make two claims about this option. First, I suggest that much of the use of the language of programming in biology is intended to gesture towards the idea of evolutionary design, without adding anything substantive about *how* the processes are carried out. To then treat the idea of genetic programming as the basis for a treatment of innateness is, in effect, to claim that being innate is the same as being an adaptation (or something close to this). But secondly, it may be possible to isolate a more substantive way in which processes of gene expression have an empirical analogy with paradigm cases of computation. The problem is that this analogy is, in a way familiar from the previous section, restricted to processes at a very low level (the level of cells and maybe small collections of cells). I then resist the attempt to "scale up" a computational description justified on this basis, to characterizes processes responsible for whole-organism phenotypes. This is a *possible* way of talking, but (again) it is justified on the basis of little more than a vague appeal to the idea of evolutionary design, and perhaps a desire to *somehow* hang onto the tempting idea that genes program the development of the whole organism.

I now discuss each of these claims in more detail. First, the "program" concept is often applied in biology in a broad and unconstrained way, guided only by vague analogies with computers and their workings. Talk of programming seems merely aimed at referring to the intricate but orderly and well-coordinated nature of many processes in biological systems. If the idea of programming adds anything, it adds a gesture towards the role of natural selection and evolutionary design.

An example of this loose way of talking about programs in biology is talk of "programmed cell death" in neuroscience. This is a process within neural development in which an excess of neurons is produced in early development and then trimmed back, to yield a functional neural architecture. As far as I can tell, the invocation of a "program" here does not indicate any significant analogy with computers, and the process could just as accurately be described as "orderly and adaptive cell death in accordance with evolutionary design."

If the idea of programming in this first sense was used to analyze innateness, then the concept of an innate trait is being treated as more or less equivalent to the concept of an adaptation – a trait that is the result of natural selection for a specific effect. This

might be qualified a little: maybe the "programmed" traits would also include some spandrel-like byproducts of directly selected traits as well, as these might be said to be indirect products of evolutionary design, though even that is questionable.

So if *being an adaptation*, or *appearing in accordance with evolutionary design*, is doing the real work here, does it do the work well? One can imagine an immediate objection that highly maladaptive, harmful traits can be innate, just as much as useful ones, and this may include traits that are not persisting via "hitch-hiking" with something more useful. Perhaps this objection can be fended off, at least by some nativists of modern stripe, with the claim that all the cases of central interest are indeed adaptations (Barkow, Cosmides, and Tooby 1992). But I think there is a deeper objection as well.

We are now drawing directly on the idea of *being an adaptation*. (I bracket the spandrels for the rest of this section.) The concept of *being an adaptation* has its own role in evolutionary theory, and that role does not line up well with the the idea of innateness as normally understood. A trait that is an adaptation must be (or rather, must have once been) heritable. That is, it must have been such that parents and offspring resembled each other with respect to the trait more than randomly chosen members of the population did (Lewontin 1970). For simplicity's sake, suppose the trait is still heritable in the present (though this will very often not be the case). This parent-offspring correlation might have a simple genetic basis, but it might not. For example, a trait can be heritable if it is socially learned by imitating a parent; this can generate parent-offspring correlation of the relevant kind. That is an extreme case, but traits can also be heritable if their appearance involves a mix of genetic and non-genetic causes, including information-using processes that are usually seen as antithetical to the idea of innateness (see also Mameli's chapter in this volume). Learning need not interfere with heritability if it is social in character, or if its contribution is stabilized by local environmental constancies. So something can be an adaptation while having a mode of development in each individual that is quite some distance from what nativists usually have in mind.

More importantly, if the appeal to "programming" merely functions to hand off the problem of innateness to the concept of adaptation or evolutionary design, then this idea should be assessed in its own right.

Secondly, however, we might be able to recognize a more substantial sense in which some biological processes are program-like. This sense would be driven by what seems to be a surprisingly close analogy between certain low-level biological process and the low-level operation of modern computers. One important category of causal process within cells is made up by cascades of up-regulation and down-regulation in genetic networks. One gene will make a product that binds to and hence down-regulates another gene, which is then prevented from making a product that up-regulates another... and so on. What we have here is a cascade of events that can sometimes be described in terms of Boolean relationships between variables. One event might only follow from the conjunction of another two, or from a disjunction of them. Down-regulation is a kind of negation, and there can be double and triple negations in a network. Gene regulation networks have a rich enough structure of this kind for it to make sense to think of them as engaged in a kind of computation. Computer "and-gates," neural "and-gates," and genetic "and-gates" have some real similarities. Most other biological processes, though just as much the product of evolutionary design, do not have a structure that motivates this sort of computational description. And once again we find, as in the case of "genetic coding," that the domain in which this computational language is well-motivated, when applied to genes, is confined to the cellular level.

So the production of individual proteins, and collections of them within a cell, might reasonably be treated as arising via a process with computational structure. As in the previous section, though, I deny that this implies a computational description of development at the whole organism level. The point is subtle, and I am aware that it might be contested. But perhaps the right way to put it is to deny that there is an *unequivocal* way to scale-up the description of cell-level processes into a description of whole-organism development in computational terms. Each of the cells in an organism has some computer-like activities. And they all act together to produce, in conjunction with environmental inputs, the whole adult organism. But the only sense in which those separate computer-like processes sum to yield a "program" at the whole-organism level is a weak sense in which we are just gesturing towards evolutionary design.

I will make one more speculative remark about this option, which bears also on the discussion of teleosemantic ideas about information. In his treatment of the concept of

innateness itself, Griffiths suggests that our thinking in nativist terms is guided by essentialist "folk biological" habits. This is an interesting hypothesis, which does draw on a body of psychological evidence, though certainly in a very tendentious way. A similar possibility might be raised, on the basis of different psychological evidence, in the case of thinking in terms of function and evolutionary design.

Rozenblit and Keil (2002) describe a phenomenon they call an "Illusion of Explanatory Depth" that is seen in people's judgments about their understanding of mechanisms. People tend to over-estimate their understanding of the structure and physical workings of various kinds of objects. Rozenblit and Keil raise various hypotheses about how and why this illusion might arise, emphasizing cases where an object has many visible parts. They suggest that the availability of a clear inventory of parts creates over-confidence about their workings. Lombrozo (2006), however, connects this work to other findings, and emphasizes a possible mechanism that Rozenblit and Keil mention briefly, but do not systematically explore. Lombrozo suggests that the phenomenon may be especially marked in cases where devices have clear functions (such as familiar artifacts), and that the illusion may come in part from the psychological effects of a sense of functional understanding. She links this idea to findings showing that children treat functional information as an inappropriate stand-in for mechanistic information; for example, they offer answers to "why" questions when asked a "how" question (Abrams, Southerland, and Cummins 2001).

So Lombrozo raises the possibility that applying a teleological description to a complex and partially unknown process might sometimes generate an over-estimation of one's mechanistic understanding; it might generate the sense that because one knows *why* one also knows *how* (or has a viable substitute for knowledge-how). I think there is a real chance that this might be operating in some discussions of teleology and gene action. The teleological gloss that is applied to gene action via talk of representation and programming adds a particular *kind* of understanding of what is going on; it gestures towards an answer to some evolutionary "why-questions." But it may also give the sense that a legitimate extraction of essential features of the causal or mechanistic facts has been achieved, when in fact this is entirely illusory.

19

5. Conclusion

The bulk of this paper has focused on the idea of making a distinction between characteristics that are coded for or programmed for by the genes, and characteristics that are not. I have argued that the defensible versions of this distinction line up badly with the idea of innateness. The defensible versions of the idea of genetic coding treat only protein molecules as coded for. The defensible versions of the idea that developmental processes involve something like computation only apply to low-level processes, so the "programmed for" traits are again too low-level to be of interest to nativists. Other versions of the idea that innate traits are "programmed to appear" are in effect a handing-off of the problem to the concept of adaptation or evolutionary design.

Where does this leave us with the concept of innateness? To me, this discussion provides some further support for the rather critical lines of analysis discussed in section 2. Options drawing directly on the idea of causation have not been exhausted, as people sometimes suppose, but the general direction of the discussion suggests that there is not any single real biological phenomenon underwriting talk of innateness. The introduction of talk of coding, programming, and the like *seems* to help with the problem, but does not. In some cases it provides no more than a misleading gloss, repackaging views that would not seem promising at all if assessed directly. If the deflationary philosophical treatments are on the right track, then at best we have in innateness a family resemblance concept which clusters a number of features together in a vague but useable way (Mameli and Bateson, newer Cowie). In some contexts, talk of innateness may function more as a disciplinary marker than anything else, removing certain phenomena from the domain of psychological explanation (Samuels, deflationary version). At worst (Griffiths), the concept is a licence for bad inferences and one that resists naturalization because of its persistent invocation of essentialist folk-biological impulses.

<p style="text-align:center">*　　*　　*</p>

References

Abrams, E., S. Southerland, and C. Cummins (2001). "The How's and Why's of Biological Change: How Learners Neglect Physical Mechanisms in their Search for Meaning." *International Journal of Science Education* 23: 1271-1281.

Ariew, A. (1996) "Innateness and Canalization." *Philosophy of Science* 63 (Proceedings): S19-S27.

Atran, S. (1990). *Cognitive Foundations of Natural History: Towards an Anthropology of Science.* Cambridge: Cambridge University Press.

Barkow, J., L. Cosmides, and J. Tooby (eds.) (1992). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York: Oxford University Press.

Block, N. (1995). "How Heritability Misleads about Race." *Cognition* 56: 99-128.

Buller, D. (ed.) (1999). *Function, Selection, and Design*. Albany: Suny Press.

Carey, S. (1985). *Conceptual Change in Childhood*. Cambridge MA: MIT Press.

Cowie, F. (1999). *What's Within? Nativism Reconsidered*. Oxford: Oxford University Press.

Cowie, F. (forthcoming). "Why Isn't Stich an ElimiNativist?" To appear in D. Murphy and M. Bishop, (eds.), *Stephen Stich and His Critics*. Blackwell.

Cullicover, P. and R. Jackendorf (2005). *Simpler Syntax.* Oxford: Oxford University Press.

Godfrey-Smith, P. (1999). "Genes and Codes: Lessons from the Philosophy of Mind?" in V. Hardcastle (ed.), *Where Biology Meets Psychology: Philosophical Essays*. Cambridge MA: MIT Press, pp. 305-331.

Godfrey-Smith, P. (2000). "On the Theoretical Role of 'Genetic Coding,'" *Philosophy of Science* 67: 26-44.

Godfrey-Smith, P. (2006). "Naturalism, Mental Representation, and Teleosemantics." In D. Papineau and G. MacDonald (eds.) *Teleosemantics*. Oxford: Oxford University Press, pp. 42-68.

Griffiths, P. (2002). "What is Innateness?" *The Monist* 85: 70-85.

Griffiths, P. and R. Gray (1994). "Developmental Systems and Evolutionary Explanation," *Journal of Philosophy* 91: 277-304.

Lewontin, R. (1970). "The Units of Selection." *Annual Review of Ecology and Systematics* 1: 1-18.

Lewontin R. (1974). "The Analysis of Variance and the Analysis of Causes." *American Journal of Human Genetics* 26: 400-411.

Lombrozo, T. (2006). *Understanding Explanation: Studies in Teleology, Simplicity, and Causal Knowledge.* PhD Dissertation, Psychology Department, Harvard University.

Mameli, M. (this volume). "Genes, Environments, and Concepts of Biological Inheritance."

Mameli, M. and P. Bateson (1996). "Innateness and the Sciences." *Biology and Philosophy* 21: 155-188.

Maynard Smith, J. (2000). "The Concept of Information in Biology." *Philosophy of Science* 67: 177-194.

Millikan, R. (1984). *Language, Thought and Other Biological Categories*. Cambridge MA: MIT Press.

Moss, L. (1992), "A Kernel of Truth? On the Reality of the Genetic Program", in D. Hull, M. Forbes and K. Okruhlik (eds.), *PSA 1992, Volume 1*. East Lansing: Philosophy of Science Association, pp. 335-348.

Moss, L. (2003). *What Genes Can't Do*. Cambridge, MA: MIT Press

Oyama, S. (1985). *The Ontogeny of Information*. Cambridge: Cambridge University Press.

Papineau, D. and G. MacDonald (eds.) (2006). *Teleosemantics*. Oxford: Oxford University Press.

Rozenblit, L., and F. Keil (2002). "The Misunderstood Limits of Folk Science: An Illusion of Explanatory Depth." *Cognitive Science* 26: 521-562.

Samuels, R. (2002). "Nativism in Cognitive Science." *Mind and Language* 17: 233-265.

Samuels, R. (this volume). "Is Innateness a Confused Concept?"

Sarkar, S. (1996). "Decoding "Coding" – Information and DNA." *BioScience* 46: 857-864.

Shannon, C. (1948). "A Mathematical Theory of Communication," *Bell Systems Technical Journal* 27: 279-423, 623-656.

Shea, N. (forthcoming). "Representation in the Genome, and Other Inheritance Systems."

Smith, B. (2002) "The Foundations of Computing." In M. Scheutz (ed.), *Computationalism: New Directions*. Cambridge MA: MIT Press.

Sober, E. (1988). "Apportioning Causal Responsibility." *Journal of Philosophy* 85: 303-318.

Stereln, K. (2000) "The "Genetic Program" Program: A Commentary on Maynard Smith on Information in Biology." *Philosophy of Science* 67: 195-201.

Sterelny, K., K. Smith and M. Dickison (1996). "The Extended Replicator." *Biology and Philosophy* 11: 377-403.

Sulloway, F. (1998). "Darwinian Virtues." *New York Review of Books* 45: 34-40.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

Wright, L. (1976). *Teleological Explanations*. Berkeley: University of California Press.