

The Use and Abuse of 'Information' in Biology

Murillo Pagnotta

Our thinking about ethical and political debates, as well as the everyday existential task of making sense of our lives, are influenced by scientific views about what genes can and cannot do and whether they determine or do not determine who we are. Consider the question of whether homosexuality (or any other characteristic, such as intelligence or body weight) results from genetic factors *or* upbringing—a question that is often put in terms of whether something is a matter of nature *or* nurture. Is that even the right question to ask? And why do we assume the disjunction in the first place?

In modern biology, the disjunction between nature and nurture is based on the idea that genes encode information about how the organism will develop—a characteristic or trait is thought of as natural if a gene is present that encodes information about its development. But the meaning of the term “information” is not as simple as it may seem, because biologists use it in different ways. It can mean the statistical correlation between a gene and a phenotype, where variation in a DNA sequence (a gene) regularly corresponds to variation in some behavior or physical characteristic of the organism (the phenotype). Or it can refer to the sense of the term developed in the mathematical theory of communication. But “information” is often used to support a stronger claim about how the genome, consisting of a collection of DNA molecules, constitutes the inherited blueprint that determines the development (and even some aspects of the behavior) of the organism.

Some of these senses of “information” are justifiable in the context of biology—for example, it is true that different alleles (genetic variants) may influence development and match with different traits, and in this sense genes have information about traits. But, as we will see, this cannot justify the privileged role commonly attributed to genes in development, nor the related privileged role of genes as carriers of hereditary information across generations, nor the reduction of evolution to changes in the

Murillo Pagnotta is a graduate student in the School of Biology at the University of St. Andrews, United Kingdom.

WINTER 2017 ~ 93

Copyright 2017. All rights reserved. See www.TheNewAtlantis.com for more information.

frequency of alleles in populations across generations. What connects and sustains these ideas is a still-dominant way of thinking about development that is indebted to preformationism—the old and discredited notion that the form of the adult organism is somehow already present in its earliest stages of development. We might say, for example, that an embryo contains in its genes a set of instructions that guides how it will develop into its own future mature form.

An alternative and more promising way of thinking about the role of genes is the developmental systems perspective (DSP). DSP is associated mostly with the work of Susan Oyama, Paul E. Griffiths, and Russell D. Gray. These theorists may disagree on specific points. But they share, to varying degrees, a dissatisfaction with the usual dichotomies of nature and nurture, genes and environment, and innate and acquired. In place of these dichotomies, they advocate relational approaches to overcome them. But before describing in more detail the DSP framework, it is worth addressing the different ways in which biologists use the concept of information and how it tends to support the idea that genes determine an organism's development.

Information About Something

Perhaps the simplest and least controversial meaning of “information” in biology is that of “covariation,” the way one thing (like a gene) varies along with another (like a trait). To explain covariation, let's choose an example from outside biology. We know that the volume of a liquid covaries with its temperature. Knowing this link allows us to use alcohol (or other liquids like mercury) to build thermometers. In this case we say that, for an observer, one variable *carries information about* the other, because knowing about one variable allows an observer to infer something about the other variable. The height of the alcohol column in the thermometer attached to my window carries information about the air temperature outside. What this means is that I, the observer, can look at the height of the alcohol column on the thermometer and infer that it is cold out there. And, by the same token, the air temperature carries information about the alcohol column. When I come back from a walk, thus knowing that the air is cold outside, I can infer more or less precisely the height of the alcohol column on the thermometer. So the values of variable A (the height of the alcohol) covary with the values of variable B (the temperature outside) because of the constraint linking them. Sometimes when two variables covary, it is because one causes the other, sometimes it is because they are

both causally linked to a third variable, and sometimes it is because there is a longer chain of causation between them.

Now let's look at an example from biology. Male humans carrying a mutation in the genes *OPN1LW* or *OPN1MW* (involved in the synthesis of the red and green photopsin proteins in the cone cells in the retina) may develop a type of red-green colorblindness. Thus a man's genotype—his genetic makeup—*carries information about* (that is, covaries with) his phenotype—his characteristics—and vice versa. If you run a genetic screening and find out that a male patient carries a mutant allele in one of those genes, you can infer that he probably has red-green colorblindness. Alternatively, if you know that he has red-green colorblindness you can infer that he probably carries a mutant allele.

Similarly, when an experimenter raises organisms with matching genotypes (that is, twins or clones) in different environmental conditions, the phenotype may covary with the environment. For example, when eggs with clones of the American alligator are incubated at different temperatures, they will develop to become either females or males, depending on the temperature. In this case, the temperature of incubation *carries information about* the sex that an observer can expect to develop in the egg. If you incubate eggs at 86 degrees Fahrenheit, you can predict that all the alligators will become female, and if you know an alligator's sex you can infer the approximate temperature it experienced during the critical period. Note that both genetic and non-genetic variables may be said to have information (in the sense of covariation) about the phenotype.

Code and Context

A more familiar, and in some ways more controversial use of the concept of information in biology is that of the "genetic code." Though the expression is sometimes loosely used to refer to the genome as such, for biologists it refers to the correlation that generally holds between the sequence of triplets of bases in the DNA—or, more precisely, in the messenger RNA (mRNA)—and the sequence of amino acids in the proteins they take part in producing. Proteins are synthesized in a complex network of chemical reactions among a great many number of enzymes, ribonucleotides, adenosine triphosphate, ribosomes, transfer RNA, and amino acids, all in a solution with the appropriate pH, salt concentration, pressure, and temperature. And of course some DNA. No DNA, no protein, that's for sure. But, just to be clear: no enzymes (or no ribonucleotides, or no ATP, or no ribosomes), no protein either! So, given the appropriate chemical context of

a cell, specific triplets of bases in the DNA correlate with specific triplets of bases in the mRNA and with specific amino acids in the protein.

The same set of predictable relations is observed in almost all living beings. For example, the triplet TAT in the transcribed portion of the DNA correlates with the triplet AUA in the mRNA, and this in turn correlates with an isoleucine residue in the corresponding protein. These reliable covariations can be represented in a table of correlations—the genetic code. The code, then, is like a condensed narrative about these molecular processes and their reliable ability to bring about the production of amino acid sequences that correspond to DNA sequences.

This means that once you know the sequence of bases in a stretch of mRNA, you can look at the table and predict the sequence of the amino acids in the related protein. So, for the observer, the sequence of mRNA bases carries information about (that is, covaries with) the sequence of amino acids. Conversely, the sequence of amino acids also *carries information about* (that is, covaries with) the sequence of mRNA (though because there are more possible triplets of mRNA than there are amino acids, the information carried by an amino acid about the corresponding mRNA is necessarily somewhat ambiguous—each amino acid can correspond to more than one mRNA triplet).

However, it is important to consider what this *doesn't* mean: it doesn't mean that the sequence of bases in the mRNA fully determines the sequence of amino acids. The chemical contexts in which the mRNA happens to find itself play a constitutive role, and so they, too, could be said to carry information about protein production, and not only the genes.

Although the genetic code is widely shared, it is not completely identical in every organism—there are a number of minor exceptions to its rules for different organisms scattered across the tree of life, and even within the same organisms. For example, in vertebrates, organelles known as mitochondria contain their own DNA, mRNA, and protein-production systems in which the mRNA triplet AUA correlates with the amino acid methionine rather than isoleucine, because the structure of the transporter RNA molecule that carries methionine is slightly different in mitochondria. That is, keeping the sequence of DNA bases constant, once you know *where* the amino acid sequence is produced, you can look at the appropriate table (the standard code or the vertebrate's mitochondrial code) and predict the structure of the protein. From this point of view, it would seem that it is the *location* of protein production that carries information about (covaries with) the sequence, while DNA is an invariant contextual condition.

And there are other wrinkles. For example, in eukaryotes (organisms with nucleated cells, which includes all animals, plants, fungi, and protists), when DNA is transcribed, the initial RNA is further transformed by chemical reactions in which some parts (called introns) are removed and others (called exons) are joined. Which parts get removed and which are kept is context-dependent, thus the same DNA may correlate with several alternative mRNAs and hence proteins. So this, too, complicates the idea that the sequence of bases in the mRNA determines the sequence of amino acids.

Furthermore, while the linear sequence of amino acids is one fundamental aspect of a protein, its function ultimately depends on its three-dimensional shape, and sometimes different protein subunits come together in particular ways to form the functional protein complex. These processes depend critically on the chemical contexts in which they occur and may involve further modifications by other enzymes, such as those that add sugar or phosphate molecules to the protein.

The point of reciting all these textbook facts is just this: the specificity in the relations between DNA and protein cannot be attributed to the linear sequence of bases in the DNA structure *independently of its chemical context*. The genetic codes do not operate in a vacuum. This is far from saying that DNA does not matter, because of course it does. This is also far from saying that all components play the same role, because of course they don't. What this does imply is that no DNA sequence has an intrinsic meaning independent of the larger chemical processes in which it participates. Hence the fact that genes covary with phenotype does not mean that they determine its development. Nor does it mean (as we will see later) that genes carry information in the sense of instructions that control the production of proteins or the development of the phenotype.

This point about genes and the larger contexts in which they participate also raises a conceptual issue—the question of how to conceive of the relations between an organism and the collection of molecules of which it is composed, between the whole and its parts. Zoom in on a digital photo of a lion and you will end up with discrete pixels that together compose the complete image. Zoom in on a living lion down to the molecular scale and what you see is not a static composition of molecules but a buzzing network of chemical transformations. Now shift your attention to the lion's furry skin and what you see is not a solid surface but a dynamically stable boundary that is continually being remade as molecules flow in and out of cells. And the animal itself is of course in motion too, as it perceives and acts in its environment.

How one should conceive of the relations among these different scales and levels of analysis—looking at molecules or body parts or the organism as a whole—is far from an easy question. But our way of thinking about this will also shape how we think about the relations between genotype and phenotype. For instance, if we think that all higher-level features of the organism, such as physical traits, are reducible to the lower-level properties of its molecules, then we might be more likely to say that genes fully determine traits. On the other hand, if we think that cells, tissues, organs, and the whole organism show emergent features—features that arise from their molecular constituents but are not reducible to them—we might be more likely to say that no single scale has primacy and that a shift from genes to higher-level features is not a shift from a causal source to its effect but a matter of finding possible correspondences.

Information as Communication

Another meaning of “information” that is at times used in biology comes from the mathematical theory of communication—an approach to understanding communication developed by the American mathematician and engineer Claude Shannon in the middle of the twentieth century. His focus was on the engineering problems in communication systems such as the telephone or the television, and biologists have come to borrow his ideas when thinking about the relation between genotype and phenotype as a form of communication in Shannon’s sense.

A communication system is composed of an information source (a person or thing that selects a state or message from a set of possible messages), a transmitter or encoder (which transforms or encodes the message into a suitable signal), the channel or medium (the processes that transmit the signal), a receiver or decoder (which transforms or decodes the signal back into a message more or less similar to the original), and the destination (which can be a person or a thing interpreting or using the message). Communication occurs when the message in the receiver covaries with the message selected by the sender, while the channel provides the conditions that make possible the correlation between the two. Thus the signal in the receiver may have information about (that is, covaries with) the signal in the transmitter, and weaker correlations between them imply noisier channel conditions. So covariation is again one sense of the term “information” in this context.

But in Shannon’s theory, “information” also means something else. The transmitter and the receiver are physical entities that can be in different

states at different times. The process of selecting a state from a set of possibilities involves a reduction of uncertainty, say reducing from a set of one hundred possible messages to one (the selected message). Shannon proposed a way to quantify this reduction and called it a measure of the *amount of information* in the message. This sense of information has nothing to do with covariation between signals, and it has nothing to do with the possible semantic content or meaning that the signal conveys to someone or something able to interpret it.

One way to reduce a set of alternatives to one is by asking a series of yes-or-no questions. The answer to each such question distinguishes a set into ever smaller groups that eventually will include only one final option. We can represent each answer by a binary digit or "bit" (1 for "yes" and 0 for "no"). And because each message is unique, each will produce a different sequence of answers, such that we can represent each message in the set by a unique combination of such bits (such as 1010001). Thus we can use the number of such binary decisions as a measure of the amount of information (in the sense of reduction of uncertainty) associated with a message. This quantity is usually called Shannon information or Shannon entropy because it is derived from the formula for entropy in thermodynamics. A message with low probability of occurrence carries a lot of information because it is unexpected, while a message with high probability of occurrence carries little information because it is not very surprising. And a message with the probability of 1 (that is, a message that is certain to occur) carries no information at all since it does not distinguish among alternatives, and so there is no surprise at all in its outcome.

But engineers are less concerned with particular messages than with the average amount of information generated by the source and received by the receiver, and especially with how they are coupled. The average amount of information produced in the source (or in the receiver) depends on the probability that each possible message is actually selected by the source (or receiver) and the amount of information each message contains. Using the formula for Shannon entropy, one can then calculate, for example, how much information in the receiver correlates with information in the source and how much doesn't. For example, if you are talking to me on the phone, the electric signals generated in your device (the transmitter) correlate with the electric signals and thus the sounds generated in my device (the receiver) as a result of the constraints linking them (the channel conditions more or less affected by sources of noise). We can then say that some amount of information (in the sense of Shannon entropy) was transmitted between our phones.

What has all this got to do with information in biology? As the philosopher Fred Dretske explains, “Any situation may be taken, in isolation, as a generator of information. Once something happens, we can take what did happen as a reduction of what *could* have happened to what *did* happen and obtain an appropriate measure of the amount of information associated with the result.” Using this broad application of information theory, we might describe the process of protein synthesis or development metaphorically as a process of communication. The DNA (or mRNA, it doesn’t matter for the analogy) is the transmitter, the protein or phenotype is the receiver, and all non-genetic factors that take part in the constructive process constitute the channel conditions, including those factors usually lumped together under “the environment.” Holding the environment constant, different genotypes may correlate with different phenotypes and thus we might say that some amount of information (in the sense of Shannon entropy) was transmitted from genotype to phenotype in development.

One very influential use of the metaphor of communication or information flow is in what Francis Crick, renowned for his role in establishing the molecular structure of DNA, called “The Central Dogma” of molecular biology. This “dogma” states that “the transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible.” Or, to give a more recent example, in a widely used molecular biology textbook we read that “The flow of genetic information in cells is...from DNA to RNA to protein.” This description of chemical processes in terms of information flowing between molecules resembles the process of communication of signals from a source to a receiver.

However, there are two important qualifications for this use of the term information in biology. First, distinguishing between source and channel conditions in a system of causal relations is a matter of convention. That is, what we decide to count as source is just one among the many interacting components whose changes we are currently interested in studying while grouping the others together as channel conditions. One could also hold the genotype constant and vary the environmental conditions of development, as geneticists routinely do, or as studies of identical twins to some extent do through statistical procedures. We could then conceive the environment to be the transmitter, and all other factors to be the channel conditions, including DNA. Holding the genotype constant, different environments may correlate with different phenotypes and thus we might say that some amount of information was transmitted from environment to phenotype in development. Any factor, genetic

or non-genetic, that plays a role in development is a legitimate source of information in this sense.

Second, we have to keep in mind that what is “transmitted” in this sense is a quantification of entropy and not any instruction, biological form, or other kind of meaning inherent in the genotype. But some writers readily switch from these senses of information as entropy and covariation to a sense of information as instruction, and that is when things can get even more controversial.

Genome as Instruction Manual

In the same textbook on molecular biology quoted above we also read that “the parent organism hands down information specifying, in extraordinary detail, the characteristics that the offspring shall have,” and that “not only must a cell use raw materials to create a network of catalyzed reactions, it must do so according to an elaborate set of instructions encoded in the hereditary information.” Passages such as these suggest more than that differences in the genes correlate with differences in traits (information in the sense of covariation) or that genes can be viewed as reducing uncertainty (information in Shannon’s sense). The picture here is of a manufacturer, the cell, following step-by-step instructions to create traits—information in the sense of specifications that determine the organism’s form.

One might suggest that these textbook passages are just references to the statistical correlations between genotype and phenotype. But this doesn’t seem to capture the meaning these writers have in mind. Recall that thinking about information as covariation applies also when we say that a non-genetic factor gives us information about a phenotypic trait, as in the example of covariation between hatching temperature and sex in the American alligator. It would be silly to say that different temperatures encode different instructions for making sex organs in reptiles. Yet, when talking about genes, such claims are taken as basic facts in textbooks.

Or one might suggest that this kind of talk is a reference to the specificity of the “genetic code” discussed above. When learning about protein synthesis, a student may well be instructed by the teacher on how to read the table of correlations and, given a sequence of mRNA bases, find the expected sequence of amino acids. The student may even learn to think of those correlations as written rules. The “code” then represents the constructive interactions among many cell components that lead to stable patterns of relations among DNA, mRNA, and protein. But to say that

these patterns are instructions—or specifications, or programs—written in the genome itself is not only unnecessary to convey the idea of specificity, it is also empirically imprecise, because it ignores the role that chemical context plays.

Talk of genetic information as a set of instructions implies a semantic notion of information. Genes are said to have meaningful *content*—instructions, programs, hypotheses, algorithms, recipes, blueprints, and so on—which is the primary source of form imposed through development on the raw formless materials. But the idea that genes play this special role in development is more like a premise than a conclusion. It is a contemporary version of preformationism—the outdated and mistaken idea that egg or sperm cells already contain tiny versions of a fully formed organism.

Of course, no one today claims that a tiny organism is in fact present in the egg or sperm. In contemporary parlance it is genetic information (here in its semantic sense) that is said to be present in the organism before the actual traits develop and that plays the lead role in development, while the other developmental influences play only secondary, supportive roles. As one biologist writes, “The information required to make a complete organism is contained within the genes of the genome. However, the genes alone are functionless; they need a complicated machinery of transcription and translation that is itself encoded in the genome.”

But more often biologists will say that the development of the phenotype is controlled by both the genes and the environment, or by the interaction between them. By saying this, they may feel shielded against accusations of genetic or environmental determinism. For example, in a respected textbook on evolutionary biology we read:

Because most phenotypic characteristics are influenced by both genes and environment, it is fallacious to say that a characteristic is either “genetic” or “environmental.” It is meaningful only to ask whether the differences among individuals are attributable more to genetic differences or to environmental differences, recognizing that both may contribute to the variation.

This seems entirely reasonable as long as we do not confuse the explanation of *variation* in characteristics among individuals with the explanation of the *production* of the characteristics. But note the use of the little word “most.” It seems to imply that at least *some* characteristics might not require “both genes and environment” to develop. Would these then be entirely genetically determined? Another passage of the same textbook

defines phenotype as “the physical manifestation of a genotype.” It is difficult to reconcile the idea that both genes and environment are necessary for the development of the phenotype with the idea that the phenotype is somehow already *in* the genotype, just waiting to be revealed.

Implicit in this line of thinking is the idea that there are two contrasting sources of causal control: the genome inside and the environment outside. Accordingly, there are also two types of developmental processes, both of which involve information in the sense of semantic meaning or instruction. One process reveals the information in the genome, and we have come to think of this as “nature”: the innate, the species-typical, the instinctual, the biological, or the features that are more under genetic than environmental control. The other process incorporates information from the environment, and this we have come to think of as the domain of “nurture”: the acquired, the accidental variation, the learned behavior, the cultural, or the features that are more under environmental than genetic control.

This common view of genes and environment interacting in the development of the organism leaves unquestioned the basic idea of preformationism—that the form of the organism already exists, prior to the organism’s growth and development, as information—which is why discussions about whether a given trait is the result of nature or of nurture, or how much of each, continue to thrive. As Susan Oyama explains in *The Ontogeny of Information* (2000),

Most solutions to the puzzle of how form arises, ... including the most recent biological dogma, have incorporated the assumption that form is to be explained by pointing to a prior instance of that very form. To the extent that this is true, they are of limited value in answering questions about origins and development.

Construction, not Instruction

The developmental systems perspective offers an alternative that involves several conceptual shifts. DSP thinks of cause and effect not as a linear relation but as a network of relations. It thinks of control over development not as localized in the genes (possibly complemented by some control in the environment) but as emerging in the relations among all influences involved in form-generating processes. And it thinks of form as a description of the phenotype itself and not a reference to some prior instruction for its development. While the distinction between genetic and environmental factors can be helpful to some extent and in certain

contexts, DSP invites us to consider the complexity of development in a way that cannot be captured by that distinction alone. The term “developmental system” tries to convey that complexity, as it refers to the phenotype together with the whole set of influences or resources that take part in its development throughout the entire lifespan.

Any lifespan starts with a living organism (which can be a single cell) that is from its beginning already embedded in ecological and possibly social relations—with its parents and other members of its species, symbionts, predators, and so on. To acknowledge this, DSP conceives development as a constructive process contingent on all these relations rather than as a process of revelation of the form already present as instructions in the genotype. The American developmental psychologist Gilbert Gottlieb suggested for this process the term “probabilistic epigenesis”—epigenesis being the biologist’s term of art for an organism’s development over time (including under influences beyond genes). Gottlieb wanted to capture the idea of development as the result of “bidirectional influences within and between levels of analysis,” which he listed as “genetic activity, neural activity, behavior, the physical, social, and cultural influences of the external environment.” Development is probabilistic, he explained, because the coordination between these various influences is imperfect. He contrasted this idea with “predetermined epigenesis,” the dominant view that developmental control is unidirectional, from genes, “pictured as the unmoved movers of development,” to physiological and behavioral traits.

The genome is certainly one source of developmental influence that is present from the start in the organism. It is one among many others, some internal and some external. And the organism’s initial set of traits undergoes a continuous history of transformations through a sequence of related states, as these multiple influences constrain and are constrained by each other and affect the moving form of the organism throughout its development. The organism is thus a history of becoming, and any description of a phenotypic trait is a condensed narrative of this history.

Heredity and Evolution

In DSP terms, what is inherited or passed on from one generation to the next is neither biological form nor a specification for it, but rather the means for developing it. It is the availability of similar developmental influences that accounts for the empirical observation that offspring resemble their parents. Some developmental resources are present from the start, such as the entire molecular structure of the zygote along with

molecules and structures surrounding it (for example, the maternal reproductive system in the case of mammals or the body of the host in the case of a parasite). Others may be produced later by the organism as it grows and differentiates. Still others may be incorporated from the outside or result from relations with other organisms and non-living features of the environment.

Thus in DSP the notion of inheritance is expanded to include non-genetic resources that are made available from one generation to the next, although authors may disagree over which resources should be considered inherited. Arguing for a more inclusive view of inheritance, Oyama suggests that

defining heredity as the passing on of all developmental conditions, in whatever manner, is preferable to defining it by genetic information. This does not require any distinction between hereditary and acquired traits, or even between mostly hereditary and mostly acquired ones; all it requires is some degree of association of developmental influences.

The difference, in this view, between a trait that does not reappear in later generations and a trait that does is a difference in how reliable this association of developmental influences is: the more reliable the association, the greater the likelihood the trait will reappear.

This affects how we think about evolution. Evolution by natural selection is often described metaphorically as the environment posing a problem to the species. As evolutionary biologist Richard Lewontin has argued, this implies a view of an environment that can be specified independently of the organism, and conversely of an organism that exists prior to encountering its environment. But this is never the case, since organism and environment are always in a mutual relation. At any moment in its lifespan, the biological structure of the organism determines what features of the physical world it can respond to and be affected by, and thus what constitutes its specific environment. Its structure also embodies past organism–environment relations from its own earlier development. And it may to some extent choose where to go and by doing this alter its environment through its metabolism and behavior. The organism, that is, plays an active role in determining the environmental conditions in which it lives. These organism–environment relations are primary, not secondary, for understanding its development and can have profound impact over evolutionary time, as advocates of “niche construction” argue.

But evolutionary biologists don't focus on the way individual organisms grow, develop, and live but, as one textbook puts it, on changes “*in*

the properties of populations of organisms, or groups of such populations, over the course of generations.” The textbook later notes that “because evolutionary changes in morphology [the forms of organisms] result from evolutionary changes in development, a full understanding of evolution requires knowledge of these processes.” So far so good. But since most evolutionists are wont to start with the premise that DNA controls (most of) development, it is hardly surprising that the textbook characterizes evolutionary changes in populations as “those that are ‘heritable’ via the genetic material from one generation to the next.”

If instead we start with the alternative premise of development as constructive interaction, then evolution might be redefined as “change over time in the composition of populations of developmental systems,” as Paul Griffiths and Russell Gray write in a book of essays on the developmental systems perspective on evolution. In this perspective, the study of trans-generational changes in the frequency of specific genes will continue to be important, but it will no longer be the entire story, as it is in the dominant view of evolution.

Simple and Complex

Information is an important and complex concept in modern biology, one that has several well-supported uses. Saying that genes or environmental features contain information about some phenotypic feature, for example a physical trait, is a convenient way to report or suggest a statistical relation between these variables, and to acknowledge that different resources contribute to development. It may also be a way of avoiding telling a comprehensive story of how the phenotype came to be, maybe because that story is for the most part unknown.

However, for many biologists such info-talk implies an underlying adherence to the idea that genes play a privileged role in development as the primary source of causal control (possibly complemented with some control by environmental factors). This is evident especially when people use the term “information” with semantic connotations, such as when information in the DNA is said to encode instructions, programs, or specifications. This way of talking about information is inadequate for explaining the rich interplay of influences contributing to an organism’s development—a complexity that the developmental systems perspective is better equipped to explain.

Some may fear that thinking in terms of developmental systems might lead to a kind of paralysis, because it would seem to require that in every

investigation, however local and specific, we constantly acknowledge and include multiple threads of causal networks, to the point that we always need to talk about everything. But that is not the case. No one can study or talk about an entire developmental system. Distinguishing between what is in focus from what is kept in the background, and making simplifying models, are part of business as usual in science.

To return to the initial example of homosexuality, we may indeed find statistical correlations between certain genetic markers and sexual attraction in given samples of the population (although there is little evidence so far for a “gay gene”). But jumping from claims about covariation between these two factors to claims about biological determination—saying that it is the result of nature, not nurture—is not justified, because the effects of genes cannot be determined independently of the many chemical, and therefore developmental, processes in which they participate. By the same token, our social environment also influences our emotions toward other people and our ideas about romantic relationships. But jumping from claims about social influences to claims about environmental determination—saying that it is the result of nurture, not nature—is equally unjustified, because the effects of social influences are also dependent on the developmental contexts in which they occur. Both views share a commitment to the distinction between nature and nurture that is so problematic.

The fact that we frame questions about human development and traits as a matter of nature or nurture may reveal more about established habits of thought than about how humans in all their variations actually develop. When we set aside the simple disjunctions between nature and nurture, genes and environment, innate and acquired—and a view of development focused on the idea that it follows a set of information—we begin to see what makes a living organism the extraordinary wonder that it is.