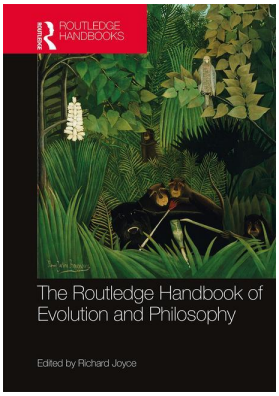

The Routledge Handbook of Evolution and Philosophy

Richard Joyce

Genetic, Epigenetic, and Exogenetic Information

Publication details
https://www.routledgehandbooks.com/doi/10.4324/9781315764863.ch8
Karola Stotz, Paul Griffiths
Published online on: 16 Aug 2017

8

Genetic, Epigenetic, and Exogenetic Information

Karola Stotz and Paul Griffiths

INTRODUCTION

The most popular account of genetic information in contemporary philosophy is "teleosemantics" (Millikan 1984; Maynard Smith 2000; Shea 2007; Kingsbury, Chapter 20 this volume). This yields a semantic notion of information—the information in a gene is a description or an instruction and as such genetic information can be true or false, obeyed or disobeyed. It also defines the information content of a gene in terms of the evolutionary history of that gene. Physically identical genes can have entirely different information content if they evolved due to different selective pressures. This way of thinking about genetic information corresponds closely to the image of genes in popular science. Genes are coded messages instructing organisms to develop in one way or another, and those instructions were written by evolution.

But the teleosemantic approach fails to meet two obvious desiderata for an account of genetic information. First, genetic information ought to feature in causal explanations of development. But the historical nature of teleosemantic information means that the information content of a gene can be changed or removed altogether without any effect on how the organism develops (Griffiths 2013): teleosemantic information is causally inert. Second, heredity is usually thought to be a precondition of evolution by natural selection, and heredity is widely supposed to involve the transmission of genetic information. Teleosemantics, in contrast, implies that the evolution of any novel character begins without any genetic information about that character being present. The genes that influence a character only carry information about it *after* that character has evolved by natural selection acting on those genes. We return to teleosemantics at the end of the chapter and discuss how it may play a role in "ultimate biology" that complements other notions of genetic information in "proximate biology." For now, however, our focus will be on a strictly proximate sense of genetic information—a sense in which that information plays a substantive causal role in the operation of living systems.

GENETIC INFORMATION: CRICK AND INFORMATIONAL SPECIFICITY

If there is one thing all philosophers who have written about genetic information agree on it is that genetic information is not merely the application of Claude Shannon's information theory to biological systems. Many components of any physical system will contain mutual information about one another merely in virtue of the fact that they form part of a causal network. But the fact that genes are "informational molecules" is meant to be distinctive, something that marks systems with a genome out from other chemical systems and also marks genes out from many other components of living systems (Maynard Smith 2000; Godfrey-Smith 2000; Bergstrom & Rosvall 2009; Stegmann 2014). In the philosophical literature, claiming that genetic information is merely Shannon information has been a way to minimize its theoretical importance (Sterelny & Griffiths 1999; Griffiths 2001).

In our view there *is* something distinctive about genetic information, but it is a distinctive property best captured using the Shannon formalism. It is a property first clearly identified by Francis Crick in the period of the mid-20th century when, it is generally agreed, molecular biology became an informational science (Kay 2000).

Although Watson and Crick had used the term "genetic information" earlier, the first substantial use is in Crick's 1958 statement of the Sequence Hypothesis and Central Dogma, some of the most influential ideas in the history of molecular biology:

> **The Sequence Hypothesis**
> In its simplest form it assumes that the specificity of a piece of nucleic acid is expressed solely by the sequence of its bases, and that this sequence is a (simple) code for the amino acid sequence of a particular protein.
>
> **The Central Dogma**
> This states that once "information" has passed into protein *it cannot get out again*. In more detail, the transfer of information from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible. Information means here the *precise* determination of sequence, either of bases in the nucleic acid or of amino-acid residues in the protein.
>
> (Crick 1958: 152–153 italics in original)

Here Crick identifies the specificity of a gene for its product with the information coded in the sequence of the gene. By doing so, he links the idea of information very closely to one of the fundamental organizing concepts of biology. Biological specificity is nothing less than the "orderly patterns of metabolic and developmental reactions giving rise to the unique characteristics of the individual and of its species" (Kleinsmith 2014). From the second half of the 19th to the first half of the 20th century specificity was "the thematic thread running through all the life sciences" (Kay 2000: 41), starting with botany, bacteriology, immunology, and serology. Specificity came to be understood in terms of the complementary three-dimensional shapes of biomolecules that exhibit specificity for one another. By mid-century, quantum mechanics had provided the necessary insight to explain the observed structural complementarity between molecules in terms of the quantum-physical forces that allow biomolecules to form weak hydrogen bonds with one another.

Crick introduced a new, more abstract conception of specificity in terms of how one molecule can precisely specify the linear structure of another. The information that specifies the product is no longer carried by a three-dimensional structure but instead by the linear, one-dimensional order of elements in each sequence—their *colinearity*. Amongst other consequences, this means that specificity becomes independent of the medium in which this order is expressed (i.e., DNA, RNA, or amino acid chain) and of the kind of reaction by which the specificity is transmitted (i.e., transcription or translation). The same information/specificity flows continuously through these three media and through both processes.

According to Crick, the process of protein synthesis involves "the flow of energy, the flow of matter, and the flow of information." While he noted the importance of the "exact chemical steps," he clearly separated this flow of matter and energy from what he regarded as "the essence of the problem"—namely, the problem of how to join the amino acids in the right order. The flow of "hereditary information," defined as "the specification of the amino acid sequence of the protein," solved this critical problem of "sequentialization" (Crick 1958: 143–144). In "Central dogma of molecular biology" Crick clarified his earlier position:

> The two central concepts which had been produced . . . were those of sequential information and of defined alphabets. Neither of these steps was trivial. . . . This temporarily reduced the central problem from a three dimensional one to a one dimensional one. . . . The principal problem could then be stated as the formulation of the general rules for information transfer from one polymer with a defined alphabet to another. (Crick 1970: 561)

The philosopher Gregory Morgan[1] corresponded with Crick late in his career about the original inspiration for using the term "information." This 1998 correspondence shows the consistency of Crick's view over forty years. He states that "information" was "merely a convenient shorthand for the underlying causal effect"—namely, the "precise determination of sequence." Information for him meant only "detailed residue-by-residue determination."

Amongst the many virtues of Crick's conception of information as the encoding of specificity through the precise determination of the sequence of gene products is that it can be made precise using information-theoretic tools. In the next section we show how this is done. In the subsequent sections we show that the resulting measure of "Crick information" can be applied to a wider range of components of biological systems than Crick himself supposed.

INFORMATION, BIOLOGICAL SPECIFICITY, AND CAUSAL SPECIFICITY

Rather than rest with an intuitive notion of specificity we can make the idea precise. Sahotra Sarkar proposed that one variable is a biologically specific cause of another if there is a bijective mapping between the values of the variables. Each value of the first variable corresponds to one and only one value of the other variable, and vice versa (Sarkar 2005: 267).

This analysis of biological specificity is identical to the analysis of a broader notion, "causal specificity," given independently by James Woodward (2010).

Woodward's account of causal specificity is part of a larger program to analyze the idea of causation as it figures in the practice of the special sciences, including biology (Woodward 2003, 2014). According to this "interventionist" or "manipulationist" view of causation, two variables are causally related if it is possible to manipulate the value of one by intervening on the other. In the limiting case, variables are causally related if there is a single pair of values of each variable that are related in this way, even if the two are unrelated across the rest of their ranges. Clearly, then, the interventionist theory of causation needs to differentiate between causes in various ways—to identify causes that "are likely to be more useful for many purposes associated with manipulation and control" (Woodward 2010: 315). A number of different ways to distinguish types of causes have been suggested, and one of these is causal specificity.

The intuitive idea is that interventions on a highly specific causal variable C can be used to produce any one of a large number of values of an effect variable E, providing what Woodward terms "fine-grained influence" over the effect variable (Woodward 2010: 302). In earlier work we and our collaborators have developed an information-theoretic framework in which to measure the specificity of causal relationships within the interventionist account (Griffiths et al. 2015). Our proposal formalizes the simple idea that the more specific the relationship between a cause variable and an effect variable, the more information we will have about the effect after we perform an intervention on the cause. This led us to propose a simple measure of specificity:

> *SPEC: the specificity of a causal variable is obtained by measuring how much mutual information interventions on that variable carry about the effect variable*

Specificity measures the mutual information between *interventions on* C and the variable E. This is not a symmetrical measure because the fact that interventions on C change E does not imply that interventions on E will change C. Formally, $I(\hat{C}; E) \neq I(\hat{E};C)$, where I is mutual information and $\hat{C}$ is read "do C" and means that the value of C results from an intervention on C (Pearl 2009).

This measure adds precision to several aspects of the interventionist account of causation. Any two variables that satisfy the interventionist criterion of causation will show some degree of mutual information between interventions and effects. This criterion is sometimes called "minimal invariance"—there are at least two values of C such that a manipulation of C from one value to the other changes the value of E. If the relationship $C \rightarrow E$ is minimally invariant—that is, invariant under at least one intervention on C—then C has some specificity for E, that is, $I(\hat{C}; E) > 0$ (Griffiths et al. 2015).

We propose that causal relationships in biological systems can be regarded as informational when they are highly causally specific. Biological specificity, whether stereochemical or informational, seems to be simply the application of the idea of causal specificity to biological systems. The remarkable specificity of reactions in living systems that biology has sought to explain since the late 19th century can equally be described as the fact that living systems exercise "fine-grained control" over many variables within those systems. Organisms exercise fine-grained control over which

substances provoke an immune response through varying the stereochemistry of recognition sites on antibodies for antigens. They catalyze very specific reactions through varying the stereochemical affinity of enzymes for their substrates, or of receptors for their ligands. Organisms reproduce with a high degree of fidelity through the informational specificity of nucleic acids for proteins and functional RNAs. Genes are regulated in a highly specific manner across time and tissue through the regulated recruitment of trans-acting factors and the combinatorial control by these factors and the cis-acting sites to which they bind. These are all important aspects of why living systems appear to be "informed" systems, and what is distinctive about all these processes is that they are highly causally specific.

GENETICS AND EPIGENETICS: TWO SOURCES OF CAUSAL SPECIFICITY

Elsewhere we have termed the encoding of specificity for the linear sequence of biomolecules "Crick information" (Griffiths & Stotz 2013). If a cause makes a specific difference to the linear sequence, it contains Crick information for that molecule. This definition embodies the essential idea of Crick's sequence hypothesis, without in principle limiting the location of information to nucleic acid sequences as Crick does. Our definition of Crick information can be applied to other causal factors that affect the sequence of biomolecules.

Crick's Central Dogma and Sequence Hypothesis were based on a very simple picture of how the specificity of biomolecules is encoded in living cells. They assume that the sequence of the gene not only *precisely* determines the sequence of the product, but also *completely* determines it. We now know that in eukaryotes additional information is needed to regulate gene expression. The large discrepancy between the number of coding sequences and the number of gene products led to the insight that the informational specificity in coding regions of DNA must be *amplified* by other biomolecules in order to specify the whole range of products. Additional specificity of a kind not captured by the original sequence hypothesis is required.

Different mechanisms of gene regulation co-specify the final linear product of the gene in question, first by *activating* the gene so it can get transcribed, second by *selecting* a chosen subset of the entire coding sequence (e.g., alternative splicing), and third by *creating* new sequence information through the insertion, deletion, or exchange of single nucleotide letters of the RNA (e.g., RNA editing). Thus specificity, and hence Crick information, is distributed between a range of factors in addition to the original coding sequence: DNA sequences with regulatory functions, diverse gene products such as transcription, splicing, and editing factors (usually proteins), and non-coding RNAs (Stotz 2006).

Absolute specificity turns out to be not inherent in any single biomolecule in these molecular networks but induced by regulated recruitment of many molecules and combinatorial control of transcription and post-transcriptional processing by those molecules (Ptashne & Gann 2002). And it is here that we will find that the networks cannot be reduced to DNA sequences plus gene products. The recruitment, activation, or transportation of transcription, splicing, and editing factors renders them functional and allows the environment to have *specific* effects on gene expression. Not just morphogenesis at

higher levels of organization, but even the determination of the primary sequence of gene products is a process of "molecular epigenesis" that cannot be reduced to the information encoded in the genome alone (Stotz 2006; Griffiths & Stotz 2013).

Eva Jablonka and Marion Lamb were amongst the first biological theorists to see the full significance of epigenetics for understanding biological information. They wrote that "DNA is not just a passive information carrier, it is also a responsive system" (Jablonka & Lamb 1995: 2). But this insight has older roots. In an immediate response to Crick's new picture of sequential information coded in DNA, the ciliate biologist David L. Nanney pointed out that:

> This view of the nature of the genetic material . . . permits, moreover, a clearer conceptual distinction than has previously been possible between two types of cellular control systems. On the one hand, the maintenance of a "library of specificities," both expressed and unexpressed, is accomplished by a template replicating mechanism. On the other hand, auxiliary mechanisms with different principles of operation are involved in determining which specificities are to be expressed in any particular cell. . . . these two types of systems . . . will be referred to as "genetic systems" and "epigenetic systems." (Nanney 1958: 712)

What is remarkable about Nanney's discussion is the rapidity with which he realized that the new view of the gene proposed by Crick implied the necessity for epigenetic mechanisms that have since been uncovered. Nanney hypothesized that the utility of epigenetic control systems "lies precisely in their ability to respond specifically to altered environmental conditions" and suggested that the influence of these systems should be understood in terms of their "specificity of induction" of developmental effects (Nanney 1958: 713, 715). Elsewhere Nanney likened them to "signal interpreting devices, yielding predictable results in response to specific stimuli from inside and outside the cell" (Nanney 1959: 333). Very much in line with this idea, Griffiths and Stotz (2013) have argued that epigenetic factors relay environmental signals to the genome.

The existence of two sources of developmental information implies that heredity needs to provide both genetic and epigenetic information if it is to reproduce living systems. In the following sections we make some distinctions between different senses of "epigenetic" and explore the *complementary* roles of the different forms of biological information.

EPIGENETIC AND EXOGENETIC INHERITANCE

"Epigenesis" is the ancient idea that the outcomes of biological development are *created* during the process of development, not *preformed* in the inputs to development. In earlier centuries, epigenesis was contrasted to the "preformationist" view that organisms already exist in miniature within sperm or ova (Roe 1981). Evolutionary developmental biologist Brian Hall notes: "As a continuation of the concept that development unfolds and is not preformed (or ordained), epigenetics is the latest expression of epigenesis" (Hall 2011: 12).

The related term "epigenetics" was introduced by Conrad H. Waddington in a broad sense that is almost synonymous with "development" (Waddington 1940, 1942). Epigenetics was the study of the causal processes by which many genes interact with one another and with many environmental factors to produce an organism. Today, however, most biologists understand the term in a narrower sense, ultimately derived from the work of Nanney quoted above (Nanney 1958; Haig 2004). In this narrower sense, epigenetics is the study of the mechanisms that determine which genome sequences will be expressed in a cell and how they will be expressed. Epigenetic mechanisms control cell differentiation in multicellular organisms, or the life cycle of unicellular organisms. Epigenetic mechanisms give cells their identity as cells of a particular type. When epigenetic modifications are maintained through mitosis this produces cell-line heredity.

This ambiguity in the term "epigenetic" itself is relatively unproblematic, but it produces a genuinely confusing ambiguity when biologists talk of "epigenetic inheritance." In the narrow sense of "epigenetic" there is epigenetic inheritance only in cases when a methylation pattern, chromatin modification, or the like is transmitted through the germline from one generation to the next. That is to say, when the mechanisms that make cell identity mitotically heritable, between cell generations, also make some aspect of cell identity meiotically heritable, between generations of whole organisms. However, the term is often meant far more broadly, to include every mechanism by which parents can influence the phenotypes of their offspring other than through the inheritance of nuclear DNA. To avoid confusion we have suggested referring to this broader class of mechanisms as "exogenetic inheritance" (Griffiths & Stotz 2013: ch. 5). This leaves us with three categories of inheritance—genetic, epigenetic, and exogenetic—each of which is a source of information for the developing organism. Our category of exogenetic comprises both Jablonka and Lamb's (2005) behavioral and symbolic inheritance systems.

To add to the confusion, many forms of exogenetic inheritance are *mediated* by epigenetic inheritance. The epigenetic modifications in question do not pass through the germline but are reconstructed anew in each generation. Some call these "transgenerational epigenetic effects" (Youngson & Whitelaw 2008) or "experience-dependent epigenetic inheritance" (Danchin et al. 2011). Despite involving epigenetic mechanisms, these are examples of *exogenetic* heredity because the epigenetic marks are not inherited by one cell or organism from another, but reestablished via an environmental influence (often reliably produced by a parent). In one well-studied example, epigenetic mechanisms have been shown to mediate the behavioral inheritance of stress reactivity and maternal care behavior in rats (Meaney 2001). Maternal behavior in the form of licking and grooming establishes stable patterns of methylation in certain genes in the pups' hippocampus. These affect gene expression and therefore brain development, with downstream effects on the behavior of the next generation of mother rats. The behavior of these second-generation mothers reestablishes the patterns of methylation in her pups without the actual patterns of methylation being inherited via the germline. Her epigenetic gene expression pattern predisposes the new mother to recreate the behavior of her mother. Interestingly, the parent–offspring correlations created through these transgenerational epigenetic effects are often much higher than those observed in narrow-sense epigenetic inheritance, like the famous case of the agouti mouse (Wolff et al. 1998).

THE EVOLUTIONARY SIGNIFICANCE OF GENETIC, EPIGENETIC, AND EXOGENETIC INHERITANCE

Evolutionary biologist Adam S. Wilkins gives a clear statement of a conventional view about the evolutionary significance of epigenetic inheritance:

> If an epimutation is to have evolutionary importance, it must persist. . . . This matter is central to whether epimutations can be treated as equivalent to conventional mutations or whether, if they have some degree of stability, some new population genetic theory is needed. (Wilkins 2011: 391)

Some cases certainly meet this criterion. In a comprehensive review of transgenerational epigenetic inheritance, Eva Jablonka and Gal Raz (2009) conclude that epigenetic inheritance is ubiquitous, and has been shown to persist for up to three generations in humans and up to eight generations in other animal taxa. In plants, which lack comprehensive reprogramming of DNA in each generation, the stability of epigenetic inheritance can rival genetic inheritance. Many cases of true epigenetic inheritance, however, particularly in mammals, would not meet the criterion of multigenerational stability. Such cases may also disappoint with respect to their efficiency or fidelity of transmission, resulting in much lower parent–offspring correlation than genetic inheritance.

However, it is simply not correct that epigenetic change will affect evolution only if the changes themselves persist for more than one generation. In conventional quantitative genetics, the importance of genetics is that Mendelian assumptions let us work out what phenotypes (and hence fitnesses) will appear in the next generation as a function of the phenotypes in the last generation. Epigenetic and exogenetic inheritance both change this mapping from parental phenotype to offspring phenotype, and therefore affect evolution. Both epigenetic and exogenetic inheritance appear in quantitative genetics as "parental effects": correlations between parent and offspring phenotypes above and beyond correlations between parent and offspring genotypes, which are also not the result of a shared environment independently influencing both parent and offspring. It has long been understood that one-generation parental effects can substantially alter the dynamics of evolutionary models, and change which state a population will evolve to as an equilibrium (Lande & Price 1989; Wade 1998). Wilkins's argument appears to be a non sequitur.

Although all three forms of heredity have evolutionary significance, this does not mean that they have *the same* evolutionary significance. Epigenetic marks are sensitive to environmental factors in that they are first "established by transiently expressed or transiently activated factors that respond to environmental stimuli, developmental cues, or internal events" (Bonasio et al. 2010: 613). Hence epigenetic variations may, indeed, be less stable than genetic ones, because they are in principle reversible by the same mechanisms that induced them. This may make them more adaptive in variable environmental conditions than genetic variation (Jablonka & Lamb 1995; Holliday 2006). Many hypotheses about the evolutionary origins of epigenetic inheritance stress its value in spatially and temporally heterogenous environments, where it allows rapid responses to change. This kind of rapid heritable response is sometimes referred to as "transgenerational adaptive plasticity" (Sultan 2015).

All three forms of heredity provide information for development in the precise sense outlined above. But the widely held view that genetic inheritance is somehow more basic than epigenetic or exogenetic inheritance, and that nucleic acids are distinctively "informational molecules," is not without foundation. With the exception of some forms of structural heredity such as the membrane inheritance system, both epigenetic and some exogenetic inheritance can reasonably be thought of as systems for the hereditary regulation of genome expression. At the heart of these heredity systems, then, is the ability of nucleic acids to provide templates for the synthesis of biomolecules. The evolutionary breakthrough that came with nucleic acid heredity was the provision of "a sequestered molecular template used by cells to transfer specificities to subsequent (cellular) generations" (Sarkar 2005: 94). Moreover, the way in which Crick information is encoded in nucleic acid templates bears many of the hallmarks of a well-designed Shannon information channel (Bergstrom & Rosvall 2009). Nucleic acid heredity is a key innovation in the history of life that allowed the highly efficient transfer of large quantities of Crick information from one cell to the next.

The mistake made by many authors who have tried to identify the special role of nucleic acid in heredity, it seems to us, has been to focus on the relative importance of different heredity systems in enabling future evolution, rather than this foundational role of nucleic acid heredity in the history of life, a point not threatened by the parity thesis (see below).

A related error is to identify the special role of nucleic acids in heredity with the idea that a greater proportion of developmental information flows through the genetic heredity channel than through epigenetic or exogenetic channels. This idea is associated with attacks on the "parity thesis," "according to which the roles played by the many causal factors that affect development do not fall neatly into two kinds, one exclusively played by DNA elements the other exclusively played by non-DNA elements" (Griffiths & Gray 2005: 420; see also Griffiths & Knight 1998). One consequence of parity is that "Any defensible definition of information in developmental biology is equally applicable to genetic and non-genetic causal factors in development" (Griffiths 2001: 396; see also Griffiths & Gray 1994). Our approach to information in this chapter clearly meets this constraint, which was inspired by Susan Oyama's calls for "parity of reasoning" in nature/nurture disputes (Oyama 2000 and elsewhere). Critics have responded to the parity idea by admitting that non-genetic causes can *in principle* carry information, but insisting that *far more* information is carried by genes (e.g., Sterelny et al. 1996). The approach to information described above allows a quantitative assessment of this claim and reveals that it is more plausible as a claim about the sources of specificity for evolutionary change than as a claim about the sources of developmental specificity (Griffiths et al. 2015: 543–550).

In conclusion, all three heredity systems have evolutionary significance. The genetic heredity system may well have been optimized for the ability of organisms to transfer biological specificity between cells. This ability, perhaps the key innovation in the history of life, was dependent on the invention of nucleic acid-based heredity. However, the addition of epigenetic and exogenetic heredity systems amplifies that information, allowing a greater range of products to be specified by the same template resources. Epigenetic mechanisms also provide the control engineering that enables the flexible expression of those template resources during development, and in response to different environmental demands. At least some epi- and exogenetic mechanisms of inheritance have

been optimized to allow organisms to respond to environmental demands on timescales intermediate between individual development and genetic selection. In modern organisms the amount of information transmitted between the generations through epi- and exogenetic heredity may exceed the amount transmitted in the underlying nucleic acid template resources considered in isolation, although nothing important turns on which form of heredity "wins" in this comparison, as each clearly plays a significant role.

PROXIMATE AND ULTIMATE INFORMATION

Our concern in the previous sections has been with proximate information, information that does causal work in living systems. In this section we turn to ultimate information, to information defined in terms of evolutionary purposes, and its relationship to proximate information.

Teleosemantic accounts define the information in a biological object in terms of the effect that that object was designed to produce by natural selection. This is "ultimate information" because it is derived from ultimate biology—the study of why organisms evolved to their current state. The most thoroughly developed account is Nicholas Shea's theory of "inherited information" (2011).

Shea accepts the parity thesis, and argues that inherited information is found in all genetic and certain special environmental causes of development. Our concern is that the presence of inherited information, whether in genes or environments, cannot contribute to proximate explanations of biological development. If two organisms contain the same allele, one by inheritance and the other as a *de novo* mutation, then according to Shea the first allele contains inherited information and the second does not. But they will, of course, affect the organism that carries them in exactly the same way, all else being equal. The presence or absence of "inherited information" in Shea's sense makes no difference in development (Griffiths 2013, and see above). This highlights the urgency to identify the link between ultimate and proximate information.

Finding the link is made easier by the fact that Shea's formulation, unusually amongst teleosemantic theories, makes a connection with the Shannon formalism. According to Griffiths (2016), the essence of Shea's proposal is that a developmental cause contains "inherited information" if (a) manipulating that cause affects how the organism develops, (b) the causal variable contains mutual information about the environment, and (c) the whole system evolved because it was fitness enhancing by matching appropriate developmental outcomes to different environments. We can apply this proposal to a typical case of adaptive phenotypic plasticity, such as the development of defensive armor in water fleas when developing fleas are exposed to chemical traces of predators (Lüning 1992). The mechanism that produces this effect contains inherited information about the need for armor. Griffiths goes on to point out that we can remove the claim (c) about history from Shea's definition of information to get a corresponding proximal notion of information, which Griffiths terms "adaptive information." A developmental cause contains adaptive information if it contains mutual information about an environmental variable and affects development so that developmental outcomes and environments correlate in a way that enhances fitness. The relationship between "adaptive information" and "inherited information" is exactly the same as the relationship between an adaptive trait and a trait that is an

adaptation. Every adaptation was, by definition, an adaptive trait in the past, and an adaptive trait will become an adaptation in future generations if it is successful enough. The ideas of adaptiveness and adaptation are complementary and both are needed to describe the process of natural selection. In the same way, both adaptive information and inherited information are needed to describe the natural selection of information systems.

It is now possible to state the relationship between ultimate and proximate information. Adaptive information is a special case of proximate biological information, in the sense defined in earlier sections, where the presence of that information enhances the fitness of the organism. Ultimate—or "inherited"—information is proximate biological information that exists in current organisms because earlier instances of the same information led to the evolutionary success of ancestral organisms. The presence of DNA sequences that contain Crick information about the structure of biomolecules is explained by the adaptive advantages of having been able to produce those molecules in the past (pseudogenes that can no longer be transcribed might be said to carry "vestigial Crick information"). The presence of proximate information in epigenetic marks and exogenetically inherited factors is explained by the adaptive advantages of being able to regulate genome expression in the past, and of the ability to adaptively match the environment on shorter timescales, as explained above.

The mistake in much of the existing literature is to try to use teleosemantics to *define* the proximate information that is a substantive causal factor in the operation of living systems. Proximate information can—indeed it must—exist before inherited information in order to be selected and thereby get to count as inherited information. Furthermore, not all the proximate information we can measure in living organisms is also ultimate information, just as not every feature of an organism is an adaptation, but we can expect that much of it will be.

CONCLUSION

In 1958 Crick introduced the idea that biological specificity could take the form of information coded in the order of bases in amino acids. Here and elsewhere we have argued that the idea of biological information is best understood in a way inspired by his work (Griffiths & Stotz 2013). Crick's definition of information as the precise determination of the order of elements in a biomolecule can be analyzed using the Shannon formalism and the interventionist approach to causation. The presence of information in Crick's sense corresponds to a highly specific causal relationship between the order of elements in a biomolecule and some cause of that biomolecule, such as the nucleic acid from which it was transcribed or translated. Specificity can be measured as the mutual information between interventions on the cause variable and the effect variable (the order of elements in the downstream biomolecule).

This definition of information can be generalized. First, other causes of the structure of biomolecules can also be highly specific, and thus contain Crick information. Second, we can apply the same measures to other causal relationships within organisms, where the effect variable is not simply the order of elements in a biomolecule. For example, enzymes are highly specific for their substrates, and this relationship can be measured in the same way that we measure the specificity of a gene for its product. Biological

information, we propose, is the very same thing as biological specificity. Living systems are "informed systems" because they exercise a high degree of specificity over their internal processes. Crick information is a special case of biological information, just as the specificity of nucleic acids for their products is a special case of biological specificity.

Both Crick information and biological information more broadly can be inherited via genetic, epigenetic, and exogenetic mechanisms. But there is an important sense in which the genetic heredity system stands out from the others: nucleic acid heredity was the key innovation in evolution that allowed the highly efficient transfer of large quantities of Crick information from one cell to the next. Many other heredity systems exploit this system to achieve the transfer of specificity to the next cell or organism generation. But the special status of genetic heredity in this respect does not support some of the other claims about it in the philosophical literature. For example, it does not support the widely held view that genes contribute much more information to development than other factors. In modern organisms a small number of coding sequences produce orders of magnitude more transcripts through epigenetic regulation of gene expression, and do so in a highly regulated manner across space and time. When we look for the source of the specificity manifested in this regulated genome expression, much of it will be found in epigenetic and exogenetic sources. This is not to say that all three heredity systems play *the same* role in development, but the differences between them are not captured by saying that one contains all or most of the information expressed in development.

Nor is it the case that only genetic information is inherited. Epigenetic and exogenetic heredity also transmit information across generations. Because of this, all three forms of heredity have evolutionary significance: they all affect which phenotypes will be seen in the next generation as the result of the differential success of competing phenotypes in the previous generation. Therefore they all affect the dynamics of evolution, as can be seen in quantitative genetic models that incorporate these other forms of heredity. The idea that a heredity system affects the course of evolution in proportion to how stable individual hereditary marks are across evolutionary time is a non sequitur. Again, this is not to say that all three heredity systems play *the same* role in evolution. On the contrary, we have argued that they play complementary roles.

Finally, the biological information found in current organisms has an evolutionary history. It is possible to look at organisms either from the perspective of proximal biology, characterizing their structure and function, or from an ultimate perspective, characterizing the teleological function of those structures and functions. In the same way, biological information can be viewed proximally, as a causal factor in the operation of living systems, or ultimately, looking at its teleological significance.

NOTE

1. Personal communication. We are extremely grateful to Morgan for making Crick's replies of March 20 and April 3, 1998 available to us.

REFERENCES

Bergstrom, C. & Rosvall, M. 2009. "The transmission sense of information." *Biology and Philosophy* 26: 159–176.
Bonasio, R., Tu, S., & Reinberg, D. 2010. "Molecular signals of epigenetic states." *Science* 330: 612–616.

Crick, F. 1958. "On protein synthesis." *Symposia of the Society for Experimental Biology* 12: 138–163.
Crick, F. 1970. "Central dogma of molecular biology." *Nature* 227: 561–563.
Danchin, É., Charmantier, A., Champagne, F., Mesoudi, A., Pujol, B., & Blanchet, S. 2011. "Beyond DNA: Integrating inclusive inheritance into an extended theory of evolution." *Nature Reviews Genetics* 12: 475–486.
Godfrey-Smith, P. 2000. "On the theoretical role of 'genetic coding.'" *Philosophy of Science* 67: 26–44.
Griffiths, P. 2001. "Genetic information: A metaphor in search of a theory." *Philosophy of Science* 68: 394–412.
Griffiths, P. 2013. "Lehrman's dictum: Information and explanation in developmental biology." *Developmental Psychobiology* 55: 22–32.
Griffiths, P. 2016. "Proximate and ultimate information in biology." In M. Couch & J. Pfeifer (eds.), *The Philosophy of Philip Kitcher* (Oxford University Press) 74–91.
Griffiths, P. & Gray, R. 1994. "Developmental systems and evolutionary explanation." *Journal of Philosophy* 91: 277–304.
Griffiths, P. & Gray, R. 2005. "Three ways to misunderstand developmental systems theory." *Biology and Philosophy* 20: 417–425.
Griffiths, P. & Knight, R. 1998. "What is the developmentalist challenge?" *Philosophy of Science* 65: 253–258.
Griffiths, P. & Stotz, K. 2013. *Genetics and Philosophy: An Introduction* (Cambridge University Press).
Griffiths, P., Pocheville, A., Calcott, B., Stotz, K., Kim, H., & Knight, R. 2015. "Measuring causal specificity." *Philosophy of Science* 82: 529–555.
Haig, D. 2004. "The (dual) origin of epigenetics." *Cold Spring Harbor Symposia on Quantitative Biology* 69: 67–70.
Hall, B. 2011. "A brief history of the term and concept of epigenetics." In B. Hallgrimsson & B. Hall (eds.), *Epigenetics: Linking Genotype and Phenotype in Development and Evolution* (University of California Press) 9–13.
Holliday, R. 2006. "Epigenetics: A historical overview." *Epigenetics* 1: 76–80.
Jablonka, E. & Lamb, M. 1995. *Epigenetic Inheritance and Evolution: The Lamarkian Dimension* (Oxford University Press).
Jablonka, E. & Lamb, M. 2005. *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life* (MIT Press).
Jablonka, E. & Raz, G. 2009. "Transgenerational epigenetic inheritance: Prevalence, mechanisms, and implications for the study of heredity and evolution." *Quarterly Review of Biology* 84: 131–176.
Kay, L. 2000. *Who Wrote the Book of Life: A History of the Genetic Code* (Stanford University Press).
Kleinsmith, L. 2014. "Biological specificity." doi:10.1036/1097–8542.082900.
Lande, R. & Price, T. 1989. "Genetic correlations and maternal effect coefficients obtained from offspring-parent regression." *Genetics* 122: 915–922.
Lüning, J. 1992. "Phenotypic plasticity of *Daphnia pulex* in the presence of invertebrate predators: Morphological and life history responses." *Oecologia* 92: 383–390.
Maynard Smith, J. 2000. "The concept of information in biology." *Philosophy of Science* 67: 177–194.
Meaney, M. 2001. "Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations." *Annual Review of Neuroscience* 24: 1161–1192.
Millikan, R. 1984. *Language, Thought & Other Biological Categories* (MIT Press).
Nanney, D. 1958. "Epigenetic control systems." *Proceedings of the National Academy Sciences USA* 44: 712–717.
Nanney, D. 1959. "Epigenetic factors affecting mating type expression in certain ciliates." *Cold Spring Harbor Symposia on Quantitative Biology* 23: 327–335.
Oyama, S. 2000. *The Ontogeny of Information: Developmental Systems and Evolution* (Duke University Press).
Pearl, J. 2009. *Causality: Models, Reasoning, and Inference* (Cambridge University Press).
Ptashne, M. & Gann, A. 2002. *Genes and Signals* (Cold Spring Harbor Laboratory Press).
Roe, S. 1981. *Matter, Life, and Generation: Eighteenth-Century Embryology and the Haller-Wolff Debate* (Cambridge University Press).
Sarkar, S. 2005. *Molecular Models of Life: Philosophical Papers on Molecular Biology* (MIT Press).
Shea, N. 2007. "Consumers need information: Supplementing teleosemantics with an input condition." *Philosophy and Phenomenological Research* 75: 404–435.
Shea, N. 2011. "What's transmitted? Inherited information." *Biology and Philosophy* 26: 183–189.
Stegmann, U. 2014. "Causal control and genetic causation." *Noûs* 48: 450–465.

Sterelny, K. & Griffiths, P. 1999. *Sex and Death: An Introduction to the Philosophy of Biology* (University of Chicago Press).

Sterelny, K., Dickison, M., & Smith, K. 1996. "The extended replicator." *Biology and Philosophy* 11: 377–403.

Stotz, K. 2006. "Molecular epigenesis: Distributed specificity as a break in the central dogma." *History and Philosophy of the Life Sciences* 28: 527–544.

Sultan, S. 2015. *Organism and Environment: Ecological Development, Niche Construction, and Adaptation* (Oxford University Press).

Waddington, C. 1940. *Organisers and Genes* (Cambridge University Press).

Waddington, C. 1942. "The epigenotype." *Endeavour* 1: 18–20.

Wade, Michael J. 1998. "The evolutionary genetics of maternal effects." In T. Mousseau & C. Fox (eds.), *Maternal Effects as Adaptations* (Oxford University Press) 5–21.

Wilkins, A. 2011. "Epigenetic inheritance: Where does the field stand today? What do we still need to know?" In S. Gissis & E. Jablonka (eds.), *Transformations of Lamarckism: From Subtle Fluids to Molecular Biology* (MIT Press) 389–393.

Wolff, G., Kodell, R., Moore, S., & Cooney, C. 1998. "Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a Mice." *FSAB Journal* 12: 949–957.

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

Woodward, J. 2010. "Causation in biology: Stability, specificity, and the choice of levels of explanation." *Biology and Philosophy* 25: 287–318.

Woodward, J. 2014. "A functional account of causation; or, A defense of the legitimacy of causal thinking by reference to the only standard that matters—usefulness (as opposed to metaphysics or agreement with intuitive judgment)." *Philosophy of Science* 81: 691–713.

Youngson, N. & Whitelaw, E. 2008. "Transgenerational epigenetic effects." *Annual Review of Genomics and Human Genetics* 9: 233–257.