

# ON GENETIC INFORMATION AND GENETIC CODING

Peter Godfrey-Smith  
*Department of Philosophy*  
*Stanford University*

## 1. The proliferation of semantic descriptions of genes, and the reaction against it

One of the most striking developments in recent biology has been the proliferation of concepts such as *coding*, *information*, *representation* and *programming*, especially applied to genes. The idea that genes can be described as having semantic properties, as well as ordinary causal properties, has become so uncontroversial in many quarters that it now appears prominently in biology textbooks. Scott Gilbert's widely used developmental biology text, to pick just one example, tell us that "the inherited information needed for development and metabolism is encoded in the DNA sequences of the chromosomes" (Gilbert 1997, p. 5).

The concepts of *information* and *coding* are the most widely used semantic or quasi-semantic concepts in genetics. Throughout this paper I use the term "semantic" in a very general way, to capture a wide range of properties that involve meaning and representation. To *code for* something is an example of a semantic property. The concept of information is trickier, because of the role of mathematical information theory. Informational properties in the sense found in mathematical information theory are not fully semantic – for example, they do not involve a distinction between representation and misrepresentation. So occasionally I will also use the term "quasi-semantic" (as I did above) when I want to be so broad that I capture borderline cases – properties and relations that seem to have something important in common with clear cases of meaning and representation, but which are not clear cases themselves.

I said that the concepts of information and coding are the most widely used semantic or quasi-semantic concepts in genetics. As John Maynard Smith stresses in his own paper in this collection, there is now a large collection of related concepts used to describe genetic mechanisms – editing, proof-reading, synonymy, and so on – that derive from these central concepts. But in this pa-

per I will focus just on coding and information. My aim is to work out what role these concepts play in our understanding of genetic mechanisms – what theoretical weight they bear and what help (if any) they really give us in understanding biological systems.

As well as examining general questions about the theoretical role of these concepts, I will focus closely on one particular way in which they are often used. The concept of genetic coding is used to express a distinction between traits of organisms: some traits are coded for in the genes and others are not. Frank Sulloway, for example, discussing work in evolutionary psychology, claims:

[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)

A trait is heritable, in a given population, if there is a certain statistical tendency for individuals with similar genotypes to resemble each other with respect to that trait. Heritability is a subtle concept, but the complications do not matter here. The important point is that Sulloway is saying that a *statistical* association between a psychological trait and genetic factors is evidence for the hypothesis that there are genes that *code for* that psychological trait.

Does Sulloway just mean that there are genes that *cause* the trait, or play some causal role in producing the trait? No, more than that must be meant. For suppose that research had found that cooperative tendencies have very low heritability, and are associated with certain environmental conditions. In the language Sulloway is speaking, that would *not* suggest that there are environmental conditions that “code for” cooperative tendencies. According to the standard framework, both genes and environmental conditions *cause* traits, but only genes *code for* them. The concept of genetic coding is apparently meant to pick out a *difference* between the causal paths leading from genes to traits, and the causal paths leading from non-genetic factors to traits. The concept of information is sometimes used in genetics in the same kind of way, although it is sometimes also used in quite different ways (section 2 below).

One possible view of the role of semantic concepts in genetics is that they are used to express a crude “genetic determinist” position. I think this interpretation is a mistake. It is true that in many popular discussions, the idea of genetic coding is associated with the idea that genetic causation is inflexible and inevitable; that “genes are destiny.” But these views are not part of mainstream biological thinking. According to mainstream biology, the “expression” of the genetic message is a biochemical process with no special causal necessity attached to it, but a process with key differences from other processes involved in development and metabolism.

Given the philosophical questions raised by all semantic properties, given the methodological uncertainties illustrated by the Sulloway passage above, and given the lingering associations between genetic coding and genetic determinism, it is not surprising that some writers have objected to the very idea of genetic coding, and to other semantic descriptions found within genetics. Sahotra Sarkar opposes talk of coding because he thinks that as more details of biological mechanisms are discovered, the idea of genetic coding becomes less and less appropriate (1996). Advocates of “developmental systems theory” (Oyama 1985, Griffiths and Gray 1994) are suspicious of genetic coding because they are suspicious of all views of development which sharply distinguish the kinds of causal roles played by genetic and non-genetic factors. Philip Kitcher (forthcoming) has responded by claiming that the concept of genetic coding carries no explanatory weight, and is just a colorful mode of talk.

At the other end of the spectrum, some hope to give a precise analysis of the semantic properties found in genes, and to use this analysis to solve problems and clarify our knowledge in biology and philosophy. Within biology, John Maynard Smith (in this collection) and George Williams (1992) are examples of biologists who think that to understand evolution properly we need a good understanding of the special information-bearing role of genes (see also Szathmáry and Maynard Smith 1995). Within philosophy, the program of explaining how genes can have semantic properties has been undertaken by Kim Sterelny and his co-workers (Sterelny, Smith and Dickison 1996, Sterelny and Griffiths 1999), and by Wheeler and Clark (1999).

I will make three main claims in this paper. The first has to do with the concept of information. Like several other authors, I hold that we cannot analyze the special semantic properties of genes using the concept of information found in information theory. Information, in the standard sense, is not enough.

The second and third claims have to do with the concept of genetic coding. I argue that the concept of genetic coding does make a theoretical contribution to solving a specific, important problem in cell biology.

Although a specific and restricted concept of genetic coding can be defended, the idea of genetic coding has diffused out from its original theoretical context, and has insinuated itself into many other descriptions of biological processes. My third main claim is that when the concept of genetic coding is found outside its original home, it probably makes no positive contribution to our thinking about biological processes.

## **2. Information is not enough**

Biology has to do with causal processes, but biologists insist on also using semantic concepts to describe genes and what they do. This is not a problem



<http://www.springer.com/978-1-4020-0930-3>

In the Scope of Logic, Methodology and Philosophy of  
Science

Volume Two of the 11th International Congress of  
Logic, Methodology and Philosophy of Science, Cracow,  
August 1999

Gärdenfors, P.; Wolenski, J.; Kijania-Placek, K. (Eds.)

2002, VIII, 362 p., Hardcover

ISBN: 978-1-4020-0930-3