INNATENESS AND CANALIZATION

ANDRÉ ARIEW†‡

University of Arizona

Cognitive scientists often employ the notion of innateness without defining it. The issue is, how is innateness defined in biology? Some critics contend that innateness is not a legitimate concept in biology. In this paper I will argue that it is. However, neither the concept of high heritability nor the concept of flat norm of reaction (two popular accounts in the biology literature) define innateness. An adequate account is found in developmental biology. I propose that innateness is best defined in terms of C. H. Waddington's concept of canalization.

1. Introduction. It is commonplace in the cognitive science literature to ascribe the term "innate" to various behavioral capacities. For two stand-out examples, Chomsky (1988) is famous for contending that Universal Language is innate while Herrnstein and Murray (1994) are infamous for boldly declaring that IQ is innate. Critics contend that innateness is not well defined in biology and so should be dropped from cognitive science (Johnston 1988, Oyama 1985, Oyama 1988, Gray 1992, Lehrman 1953). In this paper I will argue that innateness makes good biological sense.

Where does one turn to acquire at least the rough characterizations to guide our intuitions? Commonly, people associate innateness with a process that is "in the genes" or "present at birth." But such associations, when taken out of scientific context, are at best hopelessly vague and at worst reminiscent of 18th century preformationism (Sober forthcoming a). I think a better place to prime our intuition pumps is in the ethology literature, especially in the work of Konrad Lorenz. Although Lorenz's account of innateness is flawed, his work provides a set of rough characterizations or diagnostic features that can be used as benchmarks for alternative proposals. In this paper I will follow a brief exposition of Lorenz with critiques of two proposals derived from concepts found in population genetics—heritability and norms of reaction. I will show that neither concept provides an adequate definition of innateness. Finally, I will offer my own proposal that stems from developmental biology. The rough idea is that the degree to which a trait is innate depends on the degree to which it is *canalized* in development.

2. From Ethology. Lorenz (1957) sought to provide natural selection explanations for the origins of certain adaptive behavioral capacities, called "instincts," commonly found in non-human animal populations. For instance, Lorenz observed that female mallards raised to reproductive age in exclusive company of pintail ducks show no sexual affinity for the pintail drakes. But upon seeing a male mallard for the first time, the female immediately engages in the sexual courtship behavior particular to its species. Remarkably, a mallard expresses courtship be-

†I wish to thank Alan Belasco, Paul Bloom, Denise Cummins, Richard Lewontin, Elliott Sober, and Denis Walsh for comments on earlier drafts.

‡Department of Philosophy, University of Arizona, P.O. Box 210027, Tucson, AZ 85721.

Philosophy of Science, 63 (Proceedings) pp. S19–S27. 0031-8248/96/63supp-0003$2.00

havior even if it had no opportunity to learn it. That is, mallards which are naturally or experimentally deprived (via "isolation-rearing experiments") of the opportunity to acquire courtship behavior through experience tend to develop it nonetheless. If courtship behavior is not acquired from environmental cues, where does it come from? Here, Lorenz asserts the basic dichotomy from which he defined innateness: what is not learned is innate. Accordingly, the mallard's courtship behavior is innate.

But, what does innate mean besides not learned? Influenced by the theory of natural selection, Lorenz reasoned that certain seemingly adaptive species-specific behavior that develop in isolation from environmental cues are products of natural selection.[1] As products of natural selection, such behavior is *genetically transmitted* from parents to their offspring. Herein lies Lorenz's proposal: an innate trait is one that is genetically transmitted as opposed to acquired by cultural transmission or individual learning.[2] To experimentally test for innateness, Lorenz promoted the use of isolation-rearing experiments. If an organism undergoing isolation develops the trait "normally," the trait is said to be innate.

Critics of Lorenz's account assert that the gene/environment dichotomy upon which his notion of innateness is defined is false (Lehrman 1953). No biological trait[3] or behavioral capacity develops independently of environmental factors. Development involves complex interactions among genes and between genes and environments. Even Lorenz's deprived organisms developed in some (minimal) environment. Hence, critics contend, Lorenz's explanation for the origin of the mallard's behavior is insufficient; the behavior did not develop in the mallard by genes alone.

According to some critics, it follows that innateness is not well defined in biology and the term should be eliminated from the biological lexicon. However, this eliminitivist position does not follow. Rather, what follows is that innateness should not be defined in terms of a rigid gene/environment dichotomy. Evidence that "innateness" must refer to *something* lies in what remains fairly uncontroversial about Lorenz's research. In what follows, I highlight a number of rough characteristics or diagnostic features that may serve as guidelines for a reformulated definition of innateness:

(a) *On ontogeny*: The issue for ethologists is to explain how individual organisms come to develop the traits they do. For Lorenz, innate ascriptions contribute to the determination of the significant factors that enter into the development of such traits.

(b) *Innateness as an environmentally stable trait*: Innateness seems to have something to do with what environment does not do to influence development in an individual. Evidence from isolation-rearing experiments and observations in the wild suggests that some traits develop normally in a range of environments, including impoverished and non-normal ones. In such cases, the environment does not prevent the trait from being manifested. Ethologists since Lorenz sometimes associate innateness with environmentally stable, as opposed to labile, traits (Hinde 1982, 86).

(c) *Innateness as a product of natural selection*: Lorenz was interested in applying natural selection to explain the prevalence of certain highly adap-

[1]Although Lorenz was talking exclusively about behavioral capacities, I take it that his account applies to any biological trait whatsoever.

[2]I am not sure Lorenz ever said this explicitly, but it captures his idea. See Richards 1974.

[3]When I refer to biological traits, items, characters, or features, I simply mean to refer to types of biological things like organs, tissues, behaviors, etc.

tive species-specific traits. This project seems promising even though it is an open question whether all such traits are explainable via natural selection. Ever since Lorenz, ethologists interested in the source of a trait's adaptedness sometimes call products of natural selection "innate" in contrast with traits that owe their prevalence to cultural transmission or individual learning.

I will now consider two proposals to ground innateness in concepts drawn from population genetics. What makes these proposals worth considering is that they allow us to talk about genetic factors of a trait in a way that is not falsified by the truism that phenotypes always require an interaction between genes and environment. In taking the rough characterizations presented by ethologists as a benchmark, I will demonstrate substantive flaws in both proposals.

3. High Heritability. It follows from the truism about genes and environments that to ask whether genes alone caused a particular phenotype to develop is nonsense. However, we may ask a related question: How much of the phenotypic differences in a population of organisms is explainable by genetic differences and how much is explainable by environmental differences? This is the question addressed by a heritability study in quantitative genetics. Accordingly, if the phenotypic differences in a population are disproportionately due to genetic differences among members of a population, the trait is *heritable* to that degree.

Heritability in the broad sense is defined as the proportion of phenotypic variance that is due to genetic variance. (Variance is a statistical term—s^2 = the average squared deviation of the observations from the mean.) In a simplified account, phenotypic variance (Vp) decomposes into two components, genetic variance (Vg), and environmental variance (Ve) such that heritability (H^2) is the proportion of the total phenotypic variance of a quantitative trait in a population that is due to genetic variation: $H^2 = Vg/Vp$.

It might seem as though a plausible account would define innateness in terms of high heritability: *a trait is innate if and only if it is highly heritable* (or, the degree to which a trait is innate is the degree to which it is heritable). This proposal has the virtue of showing how heritability can be measured in natural populations, though one must pay attention to the numerous complications in trying to establish such measures (for a clear and detailed description of such, see Sober forthcoming b, Block 1995, or Griffiths et al. 1993). A typical test for heritability for some quantative character in human populations, e.g. height, involves the study of identical twins separated at birth and raised in different environments.[4] Identical or monozygote twins have identical genes, so $Vg\ (mono\text{-}twin) = 0$. Thus, if twins are raised apart, phenotypic differences between twins are explained solely by environmental differences: $Vp\ (mono\text{-}twin) = Ve\ (mono\text{-}twin)$. Notice that this is true only if we grant several key assumptions about gene and gene-environment interactions (see Sober forthcoming b). Suppose that twins reared apart live in environments as varied as any two randomly selected individuals from the general population, that is, $Ve\ (mono\text{-}twin) = Ve$. Then, $Vp = Vg + Ve\ (mono\text{-}twin)$. After some combining, we get the result that the genetic variance of the population as a whole can be measured by observations of the phenotypic variance between members of the general population and phenotypic variance between the twins: $Vg = Vp - Vp\ (mono\text{-}twin)$. Heritability is large if twins are more similar to each

[4]Here I follow Sober's exposition in his forthcoming b.

other with respect to the trait in question than are two randomly-picked individuals from the general population: $H^2 = (Vp - Vp\ (mono\text{-}twin))/Vp$.

Unfortunately, the proposal to define innateness in terms of high heritability suffers from substantive problems. Taking the diagnostic features of innateness provided by ethology as our guide, we find that high heritability is neither a necessary nor sufficient condition for innateness.

I will start with sufficiency. Recall that ethologists believe that a distinguishing feature of an innate trait is its environmental stability. Innate traits tend to be expressed in a wide range of environments. If the definition in terms of high heritability is to capture these features of innateness, highly heritable traits should turn out to be environmentally stable. But this need not be the case. To see why, consider the hypothetical results of an adoption study on the heritability of IQ discussed in Griffiths et al. 1993:

	Biological parents	*Children*	*Adoptive parents*
	90	110	118
	92	112	114
	94	114	110
	96	116	120
	98	118	112
	100	120	116
Mean	95	115	115

Here, IQ scores between children and their biological parents are perfectly correlated; reading from top to bottom there is a 2 point step-up for the IQ scores of both children and biological parents. Further, the scores are not at all correlated between children and adoptive parents. Perfect correlation between children and biological parents and non-correlation between children and adoptive parents entails that $H^2 = 1$; all of the IQ variation among children is due to variation among the biological parents.[5] But: (i) the mean IQ score of children is identical to that of their adoptive parents, and (ii) the mean in both is twenty points higher than that of biological parents. This suggests that environmental conditions (manifested in the IQ scores of the adoptive parents) has some effect on a child's IQ score. So while IQ is highly heritable, it is nonetheless environmentally plastic. This example shows that high heritability is not a sufficient condition for innateness.

High heritability is also unnecessary for innateness. Consider what might be thought of as a paradigm human adaptation, the possession of opposable thumbs. The possession of opposable thumbs is so successful an adaptation that all "normal" humans possess human thumbs; it has nearly gone to fixation in human populations. Further, human thumbs are environmentally stable traits; like the structure of many limbs and digits, opposable thumbs develop in a wide range of viable environments. But, it turns out, depending on the population under analysis in the heritability estimate, the possession of thumbs is not necessarily highly heritable (Sober forthcoming b). To see why, recall that high heritability is defined as a ratio that reflects the total amount of phenotypic variation that is due to genetic variation (Vg/Vp). But, for human populations in which the possession of opposable thumbs is 100%, there is no variation. Hence for most human populations, heritability is *undefined* as the denominator in the H^2 ratio is 0. Now, suppose we

[5]See Suzuki et al. 1993 for an explanation for this. For a hint, heritability studies among relatives measure the ratio of the genetic correlation—the chance that two relatives share an identical allele—between relatives over their phenotypic correlations.

enlarge our population to include one or two individual humans lacking opposable thumbs. Now there is phenotypic variation in the population, so heritability is at least definable. But consider a case in which the individuals lacking opposable thumbs do so because their mothers took particular drugs during pregnancy that disrupted fetal development. Because the variation in the population is due to environmental differences, heritability will be *very low*. Hence on the proposal that innateness is high heritability, the possession of opposable thumbs is in this case (counter-intuitively) not innate.

To diagnose the problem with the heritability proposal further, innate ascriptions are supposed to tell us something about the development of a trait in the *individual* (see the ontogeny condition above). However, heritability measurements tell us no such thing. Heritability is a measure of the variation of traits in a *population*; it does not explain why individual members of a population have the traits they do (Sober forthcoming b). For example, if one were to determine that height in humans is 60% heritable, it follows that 60% of the differences in height one sees among humans can be associated with genetic differences among them. It does not follow that 60% of an individual's height (say, from mid-thigh up) is due to genes, the rest to environment. The heritability estimate provides no information about how genes and environment interact to express height in an individual.

4. Flat Norms of Reaction. Is there a related concept in genetics that may fare better than the high heritability account? Here we will consider the merits of defining innateness as a *flat norm of reaction* for a given genotype. A norm of reaction for a genotype is a graph of the pattern of phenotypes produced by a given genotype under a range of environmental conditions (Griffiths et al. 1993, 794). The figure depicts a hypothetical norm of reaction for body size of two varieties (genotypes) of some hypothetical organism across a range of temperature in which the varieties are raised.

The graph shows that B organisms are less sensitive to environmental changes than As are. B organisms express a consistent body size regardless of the temperature of their developmental environment. This fact is represented by the *flat* norm of reaction line. In general, an inflexible norm of reaction for a particular genotype indicates that the genotype produces the same phenotype across a range of (viable) environments. We thus have the following proposal: *a trait is innate for a genotype within a particular environmental range if and only if its norm of reaction is flat across the range of environments specified.*

Two points are worth mentioning here about the proposed definition. First, notice that innateness on this proposal is relative to particular genotypes. To see why, suppose some genotypes for blue eyes in humans exhibit flat norms of reaction across some environmental range. It does not follow that all blue eyes genotypes produce flat norms of reaction lines; it may be that some eye-color genotypes are more sensitive to environmental variation thereby expressing blue-eyes only in some environments. Second, norms of reaction are relativized to a specified range of environments. Consequently, it would be wrong to extrapolate from the figure that variety B genotypes are innate (i.e. produce flat norms of reaction) across, say, all "normal" environments for that variety. It may be that air quality, elevation, or any number of other environmental conditions would produce non-flat norms of reaction curves for the same genotypes.

The flat norm of reaction account of innateness has merit. Significantly, it captures the ethologist's intuition that innate traits are ones that are environmentally stable rather than labile. Unfortunately, problems plague the account. First, a flat norm of reaction such as the one shown in the figure depicts the pattern of *adult*

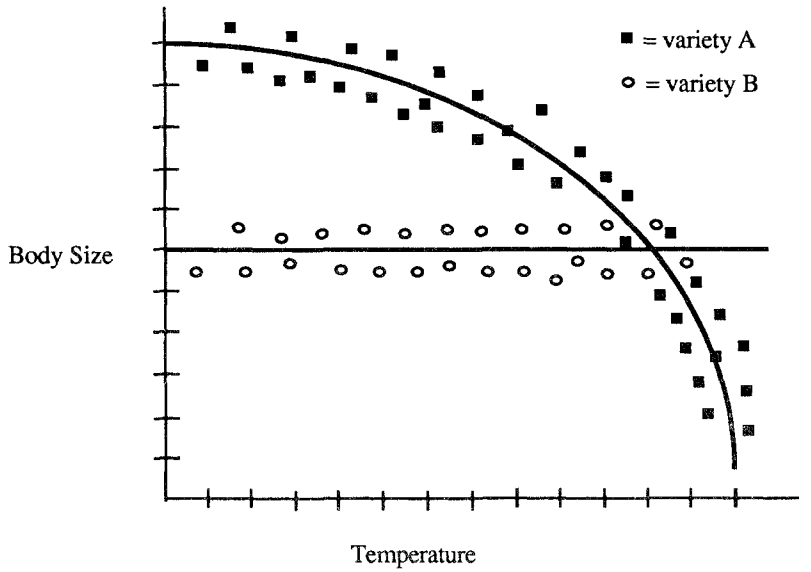

Figure 1.

phenotypes produced by a given genotype under a range of environmental conditions that are *fixed* throughout the course of development. What's missing is the possible effect a fluctuating environment *during the course of development* has on the phenotype in question. That is, what is needed is to plot *individual life histories* across a variety of environmental conditions, not norms of reactions. For example, from the figure we have no idea how fluctuating temperature throughout development affects body size for each genotype. In short, a flat norm of reaction can mislead one to conclude wrongly that a trait is stable for a genotype when in fact it is not. Therefore, flat norms of reactions are insufficient to determine innateness.

The second problem is that flat norms of reaction are properties of populations, whereas innateness is supposed to be a matter of an *individual's* ontogeny. A flat reaction norm is a measure of the *average* phenotypic scores for all individuals possessing an instance of the genotype. (More precisely, the norm of reaction is a best fitting regression line—the least square line for all individual scores measured.) Phenotypic scores of individual members of the population are represented as dots on the graph, as shown in figure. *No* individual member is represented by the regression line. Notice the scatter of dots around the regression lines. The difference between the reaction norm and the individual phenotypic score is what geneticists call "developmental noise" which indicates that there are differences in how each individual develops. The point here is that the reaction line does not represent the development of any particular member of the population; the dots do. It would be a category mistake to say that innateness *qua* property of an individual is defined as a flat reaction norm *qua* property of a population. Therefore, flat norms of reaction do not define innateness.

Nevertheless, under certain conditions a flat norm of reaction may be a good indicator or predictor of innateness in an individual just as a regression line may be a good predictor of the phenotypic score of individuals possessing an instance of the genotype. If a trait is stable in the way that we require—against fluctuations in both the initial environmental conditions and throughout the course of development—then it will produce the same phenotype no matter how the environment varies. That is to say that the trait will produce a flat norm of reaction (ideally with little or no scatter around the regression line) for the population of individuals possessing the trait. So only for some traits is a flat norm of reaction a good indicator for innateness.

Out of the various proposals and criticisms we have the makings of a new proposal to define innateness: to say that a biological character is innate is to say that an individual's development tends to express the biological item as its end-state in a wide range of initial environmental conditions and persists regardless of environmental fluctuations during the course of development. When such conditions hold, developmental biologists call this *canalization.* Let us see how the term is used in developmental biology.

5. Canalization The developmental biologist C. H. Waddington was struck by the fact that developing organisms tend to produce a number of distinct and well-defined body-types despite environmental variation. To explain the phenomenon, Waddington envisioned development as a branching out of various developmental pathways each leading to the production of a distinct end state. Once development starts in the egg, a combination of genetic and environmental factors force the developing mass down one or another pathway. For the development of some traits, once a pathway is chosen it is entrenched or bound to produce a particular end state. It is this entrenchment that Waddington called canalization. Waddington defined canalization as "the capacity to produce a particular definite end-result in spite of a certain variability both in the initial situation from which development starts and in the conditions met with during its course" (Waddington 1975, 99). The developmental pathways that lead to the development of organs and tissues are canalized such that only the most severe of environmental fluctuations can force the development of these tissues from their normal path.

For our purposes it can be said that Waddington provided an account of the development of environmentally stable traits that occur in individuals (satisfying both the ontogeny and the stability conditions above). This makes Waddington's idea a good candidate for an account of innateness: *the degree to which a biological trait is innate for a genotype is the degree to which a developmental pathway for individuals possessing an instance of that genotype is canalized.* The degree to which a developmental pathway is canalized is the degree to which it is bound to produce the end-state regardless of environmental variation in either (a) its initial state, or (b) during the course of development. Notice that this definition preserves the idea that traits are innate with respect to certain genotypes. It may turn out that, for example, some of the genotypes that typically express blue eyes are canalized while others are more sensitive to environmental fluctuations. Notice also that canalization is a matter of degree. Limb development in many organisms is highly canalized; limbs develop in all but the harshest environments. Sexual affinity in mallards may also be canalized if the results of isolation-rearing experiments are to be granted as evidence. However, a trait requiring several very specific environmental cues to develop in individuals, e.g., learning French (as opposed to learning language simpliciter), is not highly canalized.

In realizing that Waddington's notion of canalization applies well to the idea

of innateness, we open the door to a range of parallels between the work of Waddington and the work of ethologists like Lorenz. A striking example will further strengthen the case for thinking that innateness is canalization. In this instance, we find Waddington solving a problem that Lorenz failed to solve: how can we invoke natural selection to explain the origins of highly adaptive traits, such as sexual affinity in mallards, that are seemingly acquired characteristics but turn out upon further investigation (e.g., performing an isolation-rearing experiment) to manifest in isolation? For Waddington, the key is the concept of canalization.

To illustrate the phenomenon, Waddington (1975) managed to induce an extreme environmental reaction in the developing embryos of Drosophila. In response to ether vapor, a proportion of embryos expressed a radical phenotypic deviation, a second thorax. At this point in the experiment we would say that bithorax is not innate; it is a kind of chimera induced by an unusual environment. But then Waddington continually selected for Drosophila with the developmental capacity to respond to environmental stress. After about twenty generations of selection, some Drosophila were obtained that developed bithorax *without exposure to ether treatment.* What happened, according to Waddington, was that selection favored a particular pathway that led to the production of the optimal (in this case desired) effect. Eventually the pathway became canalized, hence the end-state, bithorax, appeared regardless of environmental conditions. Waddington thus showed how natural selection can, so to speak, install a trait as innate. In this case natural selection turned what was once an acquired trait—bithorax qua reaction to ether shock—into an innate, i.e., canalized, trait.

6. Conclusion. Taking its cue from critics of Lorenz, I have presented an account of innateness that avoids the fallacy of claiming that traits can develop by genetic causes alone. Innateness, on the canalization account, is a property of a developing individual. Further, the proposal captures what is thought to be distinguishing features of innateness: satisfying the ontogeny condition, referring to the capacity to produce environmentally stable traits, and making sense of the idea that natural selection can install innate traits in a natural population.

REFERENCES

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

Chomsky, N. (1988), *Language and Problems of Knowledge.* Cambridge, MA: The MIT Press.

Falconer, D. S. (1981), *Introduction to Quantitative Genetics.* 2nd edition. New York: Longman.

Gray, R. (1992), "Death of the Gene", in P. Griffiths (ed.) *Trees of Life.* Boston: Kluwer Academic Publishers, pp. 165–210.

Griffiths, A. J. F, Miller, J. H., Suzuki, D. J., Lewontin, R. C., Gelbart, W. M. (1993), *An Introduction to Genetic Analysis.* 5th edition. New York: W. H. Freeman and Company.

Herrnstein, R. J. and Murray, C. A. (1994), *The Bell Curve: Intelligence and Class Struggle in American Life.* New York: Free Press.

Hinde, R. A. (1982), *Ethology.* New York: Oxford University Press.

Johnston, T. D. (1988), "Developmental Explanation and the Ontogeny of Birdsong: Nature/Nuture Redux", *Behavioral and Brain Sciences* 11: 617–629.

Lorenz, K. (1957), "The Nature of Instincts", in C. H. Schiller (ed.), *Instinctive Behavior.* New York: International University Press, pp. 129–175.

Lehrman, D. S. (1953), "A Critique of Konrad Lorenz's Theory of Instinctive Behaviour". *The Quarterly Review of Biology* 28: 337–363.

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

———. (1988), "How Do You Transmit A Template?", *Behavioral and Brain Sciences* 11: 664–645.
Richards, R. J. (1974), "The Innate and the Learned: The Evolution of Konrad Lorenz's Theory of Instinct", *Philosophy of Social Sciences* 4: 111–133.
Sober, Elliott. (forthcoming a), "Innate Knowledge" in Routledge *Encyclopedia of Philosophy.*
———. (forthcoming b), "Separating Nature and Nuture", in D. Wasserman and R. Wachbroit (eds.).
Waddington, C. H. (1975), *The Evolution of an Evolutionist.* Ithaca: Cornell University Press, pp. 36–59.