Can Behaviors Be Adaptations?*

Catherine Driscoll†‡

Kim Sterelny and Paul Griffiths (Sterelny 1992, Sterelny and Griffiths 1999) have argued that sociobiology is unworkable because it requires that human behaviors can be adaptations; however, behaviors produced by a functionalist psychology do not meet Lewontin's quasi-independence criterion and therefore cannot be adaptations. Consequently, an evolutionary psychology—which regards psychological mechanisms as adaptations—should replace sociobiology. I address two interpretations of their argument. I argue that the strong interpretation fails because functionalist psychology need not prevent behaviors from evolving independently, and it relies on too strong an interpretation of the quasi-independence criterion. The weaker interpretation does not undermine sociobiology, and evolutionary psychology would be vulnerable to the same criticism. Finally, I offer reasons to think that both mental mechanisms and behaviors can be adaptations.

1. Introduction. Sociobiology has had a bad reputation for a long time. Some of its advocates have misappropriated the methods of the project and the glamour of "scientific" Darwinian explanations of behavior to promote conservative social values and genetic determinism. Consequently (and understandably), there has been significant motivation for scientists and philosophers troubled by these abuses to try to find ways to undermine sociobiology at large. While criticisms focused at questioning individual explanations offered within sociobiology or the practices of individual sociobiologists turn up problems (e.g., see Kitcher 1985) it is not clear that there are any substantive criticisms that apply to sociobiology as a research program.

More recently, the project of evolutionary psychology has largely replaced sociobiology as the popular form of evolutionary behavioral

*Received July 2003; revised September 2003.

†Catherine Driscoll, Department of Philosophy, Dartmouth College, 6035 Thornton Hall, Hanover, NH 03755-3592; e-mail: catherinedriscol@hotmail.com.

‡Many thanks to Stephen Stich for numerous comments on earlier drafts of this paper; to Karen Neander and Ted Sider for their very helpful remarks; and to several anonymous reviewers who made comments on an earlier submitted version.

Philosophy of Science, 71 (January 2004) pp. 16–35. 0031-8248/2004/7101-0002$10.00
Copyright 2004 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/381410 Published online by Cambridge University Press

science. It seems common to regard evolutionary psychology as more or less the same project as sociobiology, "relabeled" to help it avoid the negative connotations of the older project. However, there is a significant difference between the two. While both projects are particularly focused on finding elements of "human nature" that are *adaptations*—that is, behaviors or psychological systems that are present in the human population due to the operation of natural selection upon them[1]—the specific targets of their investigations are different. Sociobiology is primarily concerned with the origins of human social behavior,[2] whereas evolutionary psychology is interested in providing evolutionary histories for particular human psychological mechanisms[3] (Cosmides and Tooby 1987). To illustrate the difference between the projects, compare the work of the traditional-style sociobiologist Sarah Hrdy (2000) on the evolution of human maternal infanticide with the work of the evolutionary psychologists Leda Cosmides and John Tooby (1992) on cheater detection. Hrdy concentrates on describing the ecological and biological conditions operating on mothers to stimulate the evolution of infanticide behavior (and is largely agnostic about the underlying psychology). Cosmides and Tooby's experiments on the other hand, are designed to identify the cognitive mechanisms that allow humans to make judgments about whether other people are "cheating" on social contracts.

My focus in this paper will be on an argument proposed by Kim Sterelny and Paul Griffiths (Sterelny 1992, Sterelny and Griffiths 1999). It is another attempt to show that there is something fundamentally wrong with sociobiology as a whole, with the very idea of trying to identify evolutionary histories for particular units of human behavior. Sterelny and Griffiths argue that behaviors are generally not capable of being the targets of natural selection, and thus cannot be adaptations in the way that sociobiologists like Hrdy would require. They claim that the human mind is in fact a complex functional system in which most individual mental mechanisms play a role in the production of multiple behaviors, and

1. This apparent "adaptationism" in sociobiology and evolutionary psychology—i.e., their expectation that many human behavioral and cognitive traits will be adaptations—has drawn criticism from some quarters, most notably Gould (2000).

2. Sterelny and Griffiths are not clear about precisely what they take "behavior" to be; they might mean that token behaviors, e.g., a token instance of xenophobia, are the adaptations sociobiology looks for. Or, they might mean that behavioral types exhibited by individual organisms are adaptations. Or, they might mean that sociobiologists are identifying behavioral dispositions. For the purposes of this paper I will take it that they mean the latter—dispositions to behave in certain ways given specific sets of stimuli.

3. By "psychological mechanism" or "cognitive mechanism" most evolutionary psychologists mean individual components of the mind. Evolutionary psychologists regard human cognition as computational in nature and think of each individual mechanism as a kind of computer, that is, a symbol processor that executes programs or algorithms.

https://doi.org/10.1086/381410 Published online by Cambridge University Press

individual behaviors are themselves the result of the interaction of multiple mechanisms. The complex psychological interactions that produce behaviors prevent individual behaviors—for example, a conditional disposition for infanticide—from evolving independently of other behaviors; i.e., individual behaviors are not usually *mosaic*.[4] Traits that are not mosaic cannot be adaptations. Sterelny and Griffiths argue that evolutionary psychology avoids these difficulties by claiming that the underlying mechanisms, rather than the behaviors, are the adaptations. Since mechanisms can be mosaic in the appropriate way, they can be adaptations.

Sterelny and Griffiths' argument can be given two main interpretations: a strong interpretation and a weaker interpretation. The two versions of the argument face different sets of problems; however it is not clear which version of the argument Sterelny and Griffiths have in mind. In this paper I intend to argue that neither version of the argument really undermines sociobiology. The two versions depend on intuitions about the possibility of behavioral adaptations that conflict with what we know about both human psychology and general biology. Section 2 will address the strong version of the argument: I will first argue that a functionalist psychology of the sort Sterelny and Griffiths have in mind does not necessarily prevent behaviors from evolving independently of each other. Second, I will show that the notion of mosaicness that Sterelny and Griffiths are employing is too strong; and that we need a weaker notion in order for it to be possible that real physiological adaptations have evolved by natural selection. But if we employ this weaker notion, behaviors could still count as evolving independently of each other, and Sterelny and Griffiths' argument will fail. In Section 3 I will address the weaker interpretation of the argument. This form does seem to go through; however, it does not really undermine sociobiology as a project, and functions more as a caution to the practices of sociobiologists in their search for adaptationist explanations of behavior. If sociobiologists are sufficiently careful the problems Sterelny and Griffiths raise can be avoided. The weaker interpretation of Sterelny and Griffiths' argument also does not support evolutionary psychology over sociobiology, since evolutionary psychology needs to operate with the same caution as sociobiology and for similar reasons. In Section 4, having disposed of the arguments against the idea that behaviors can be adaptations, I will offer some reasons for thinking that we may need to regard both behaviors *and* psychological mechanisms as adaptations in different cases.

4. The term "mosaic" is a technical term fairly widely used in biology—usually in the context of "mosaic evolution"—where one or more traits evolve approximately independently from the other traits of an organism. I will give a clearer explanation of the term in a later section of the paper.

https://doi.org/10.1086/381410 Published online by Cambridge University Press

2. The Strong Form of Sterelny and Griffiths' Argument. A strong interpretation of Sterelny and Griffiths' argument runs roughly as follows:

1. In order for sociobiology to be a worthwhile project, it must not be impossible for most behavioral traits $B_1 \ldots B_n$ to be adaptations.
2. For any B_i, if B_i is to count as an adaptation, then B_i must be the possible target of natural selection.
3. For any B_i, in order to be the target of natural selection B_i must be *mosaic* (i.e., B_i must meet Lewontin's (1978) quasi-independence criterion[5]) (Sterelny and Griffiths 1999, 320; Sterelny, personal communication).
4. In order for B_i to be mosaic there must be a one-to-one correspondence between B_i and the mechanism M_{Bi} that produces B_i (Sterelny 1992, 168).
5. For most B_i of $B_1 \ldots B_n$ there is no one-to-one correspondence between M_{Bi} and B_i (Sterelny 1992, 168; Sterelny and Griffiths 1999, 321).
6. Therefore sociobiology is not a worthwhile project.

Most of the weight in this argument rests on Premise 4, and that is where Sterelny and Griffiths' argument is most problematic. I take premises one to four to be the most important premises of the argument. In Sections 2.1 to 2.4 I will explain why I attribute these premises to Sterelny and Griffiths and address the issues that they raise.

2.1 Premise 1: Does Sociobiology Require that Most Behaviors Could Be Adaptations? Premise 1 states that *most* behaviors must be adaptations in order for sociobiology to be worth pursuing. Sterelny claims that sociobiology is concerned with the adaptive significance of behaviors—i.e., it tries to provide evolutionary histories for behaviors, and in so doing, "aims wrongly" (1992, 168). So obviously, if Sterelny and Griffiths can provide us with some prima facie reason for thinking that most behaviors cannot be adaptations, sociobiology is no longer worthwhile, since there will not be any subject matter for it to study.

However, Sterelny and Griffiths think that their argument does not show there is a prima facie reason for thinking that no behaviors are adaptations, since some behaviors *will* have evolutionary histories of their own:

> Our own species runs a few stimulus response mechanisms isolated from the rest of what's within. But reflex behavior is a small and relatively uninteresting chunk of human behavior . . . our behavior is the result of perceptual inputs, our learning history and very complex

5. I will explain Lewontin's (1978) criterion shortly.

https://doi.org/10.1086/381410 Published online by Cambridge University Press

> interactions between distinct psychological mechanisms. (Sterelny 1992, 168)

Sterelny's point is that there are behaviors that have an evolutionary history—e.g., reflex behaviors and behaviors that result from other stimulus response mechanisms (Sterelny 1992), but that those behaviors are not the subject matter that most scientists of human behavior find interesting: e.g., incest taboos, infanticide, xenophobia. It might be possible to have a science that investigates reflexes, but it would not be a very interesting science to sociobiologists. As a consequence I take it that we should read the "most" in premise one to mean "those among $B_1 \ldots B_n$ that sociobiologists find interesting topics of study."

2.2 Premise 2: Adaptation and Natural Selection. Premise 2 is not controversial. In order for a trait to count as an adaptation, it must be possible for natural selection to act on it. The upshot is, if we want to pursue a project that recognizes certain kinds of trait as adaptations, we must not have any a priori reason to think that that sort of trait cannot be accessible to natural selection.

2.3 Premise 3: Adaptation and the Quasi-Independence Criterion. Sterelny and Griffiths' third premise invokes the notion of a "mosaic trait" that is common in biology—in order for a trait T to be an adaptation T must be *mosaic,* i.e., (roughly) T must be able to evolve independently of other traits. A trait that cannot change without significantly affecting other traits is called a *connected* trait. Sociobiology's "main failing" is that behaviors are not mosaic and thus are not adaptations (Sterelny and Griffiths 1999, 320.) Biologists usually seem to understand mosaicness in the following way: suppose we have an organism with traits $T_1 \ldots T_n$. If natural selection is to count as acting on a trait T_i (i.e., if T_i is to count as an adaptation), its effect on T_i must be distinguishable from its effect on any other trait T_k. Otherwise there is no selection for T_i. It must be distinguishable in the sense that the positive fitness benefit of the changes to T_i must outweigh the negative changes to T_k. Even if the changes to T_i confer a fitness benefit on O, that benefit can be masked if T_i is connected to T_k and the consequent changes to T_k have a negative effect on O's fitness that outweigh the positive changes to T_i.

Possibly the most famous way of expressing what it means to be a mosaic trait is Lewontin's quasi-independence criterion. Lewontin (1978) claims that in order to be targets of natural selection, traits need to be 'quasi-independent':

> . . . quasi-independence means that there is a great variety of alternative paths by which a given characteristic may change, so that

https://doi.org/10.1086/381410 Published online by Cambridge University Press

> some of them will allow selection to act on a characteristic without altering other characteristics of the organism in a countervailing fashion; pleiotropic and allometric relations must be changeable. (1978, 230)

I take it that Lewontin means there are numerous ways in which any given trait T of an organism can be changed. T is produced by a package of genes, and a set of environmental conditions, structural relationships and so forth. Changes to T occur when there are changes to any one or more of those genes, environmental conditions, or structural relationships. Changing those things may (and probably will) have other effects in the organism. In order for a trait to be the target of natural selection, there must be at least one way to change T that does not result in *countervailing* changes in the rest of the organism. Otherwise the fitness effect of T is not distinguishable from the effect of the other traits; T cannot be selected for—T cannot be an adaptation.

How strongly are we to interpret the quasi-independence criterion? The concept of quasi-independence is problematic. First, connectedness is a matter of degree. A trait T can be connected to different numbers of other traits; those connections can be easy or difficult to break genetically; those connections may or may not manifest themselves in all environments. The connected traits may be either positive or negative or neutral with respect to fitness, and vary in the degree to which they affect fitness. As a consequence, it is possible for T to be selected for whilst having a slight effect on other traits; but not an effect that is so dramatic or damaging to the organism as a whole that T's fitness benefit is masked.

The second problem with the concept of quasi-independence is that connectedness is very common. Genetic connections between traits can occur because the genes for those traits occur close together on the same chromosome, or because there is a single gene that contributed to more than one trait (pleiotropy) or because the presence of one gene affects the expression of another gene, and so forth. Some traits are connected because they are the consequence of the developmental program of the animal—an example of this would be lung number—the human left lung and right lung are connected because they are the consequences of the bilateral symmetry component of the human developmental program. Some traits may occur together because they are structurally connected—they are the consequence of growing other parts of the body in a certain way: e.g., the human chin is the consequence of the two human jaw bones meeting in the middle; or traits may affect each other because of their allometric relationships—changing how one develops forces overall changes in the shape, size and relative proportions of other parts of the organism. If connectedness is common, and if we take the quasi-independence

https://doi.org/10.1086/381410 Published online by Cambridge University Press

criterion to mean that there must be a way to evolve a trait which has *no* negative effects on other traits, then there are going to be very few traits capable of counting as adaptations on Lewontin's criterion. I do not think this is Lewontin's intention.

I interpret Lewontin's quasi-independence criterion in the following way: there are numerous ways in which any given trait T of an animal can be changed. Changing T may (and probably will) have other effects in the organism. In order for a trait to be the target of natural selection, there must be at least one way to change T that does not result in *overwhelmingly detrimental* changes in the rest of the organism. More precisely:

> (QIC) A trait T_i with connections to other traits $T_1 \ldots T_n$ can evolve so long as there is at least one way to change T_i such that the positive contribution of T_i to an organism's fitness is greater than the total negative contribution supplied by any changes to $T_1 \ldots T_n$.

The upshot of this is that T_i can be involved in significant pleiotropic and other kinds of connections and still meet the independence criterion.

Certain things that Sterelny says suggest he has a stronger understanding of the independence criterion than I do. Consider the following quote:

> If aggression, rape or xenophobia are the result of interactions amongst mechanisms, they have no biological functions, for these behavioral activities will have no Darwinian histories. Not only will xenophobic behavior have no function, there will be no specific, isolatable disposition, no specialized internal base, for that behavior. Xenophobe building is possible only if there is a complex of genetic changes and developmental pathways whose only adaptively significant result is xenophobia when, or mostly when, that is fitness enhancing. (Sterelny 1992, 168–169)

Sterelny seems to be claiming that the quasi-independence[6] criterion requires there to be *no adaptively significant result* from the genetic change that produces xenophobia other than the xenophobia if xenophobia is to count as an adaptation. If by this he means there should be no other fitness effects on human psychology as a consequence of introducing xenophobia, beyond the fitness contribution of xenophobia itself, then I disagree. All that is required is that the consequence of the genetic changes to the mechanism(s) that produce xenophobia is not so negative that it outweighs the positive fitness contribution of the xenophobia.

6. In the text Sterelny and Griffiths usually only refer to the notion of a "mosaic trait"; however in personal communication Sterelny indicated that he understands mosaicness in terms of Lewontin's criterion—which is why I employ it here.

https://doi.org/10.1086/381410 Published online by Cambridge University Press

2.4 Premise 4: Mosaic Behaviors Require One-to-One Correspondence Between Behaviors and Mechanisms. The claim on which all of Sterelny and Griffiths' argument hangs is really Premise 4. If behaviors are to be adaptations they must evolve by natural selection, and evolution by natural selection requires that behaviors be mosaic—i.e., they are able to change without causing countervailing changes in other traits. Sterelny (1992) claims that behaviors cannot evolve because there is "*nothing like* a one-one correlation between behaviors and mechanisms" (his italics, 168). Why does Sterelny think this correlation is important?

Sterelny and Griffiths' concern is that there is something about the structure of the psychology that produces human behaviors that prevents those behaviors from evolving truly independently of each other:

> . . . The right psychology for humans is a functionalist psychology. Our behavior is the result of perceptual inputs, our learning history, and very complex interactions between distinct psychological mechanisms. (Sterelny 1992, 168)

It is this complex psychology that Sterelny and Griffiths think generates the connections between behaviors that make it impossible for those behaviors to truly be evolutionary atoms—independent targets of selection. The idea is that behaviors are not mosaic in the right way because there is no one-to-one connection between behaviors and the mechanisms that produce them; each behavior is produced by a complex of mechanisms, and each mechanism in turn is involved in the production of multiple behaviors:

> The human behavioral repertoire is not an aggregation of independent units. Our behavior is produced by mental mechanisms that play a role in many different behaviors. Some of the mental mechanisms used in hunting are used in storytelling. . . . Such behaviors might be alterable only by altering the underlying mental mechanisms, and since these mechanisms are used for many different purposes, any change in them would have many other consequences. Hence individual behaviors are unlikely to have histories to call their own, or to have independent adaptive significance. (Sterelny and Griffiths 1999, 321)

Sterelny and Griffiths think that this lack of one-to-one correspondence makes the independent evolvability of behaviors all but impossible. There does seem to be a certain lack of plausibility in this claim. After all, there are examples in biology of physical functions which do not occur as a consequence of a single physical mechanism, but due to the interaction of several mechanisms. One example is digestion—this is a function that is produced by the interaction of stomach and intestinal movement, the

https://doi.org/10.1086/381410 Published online by Cambridge University Press

excretion of enzymes by various organs, and a blood supply that takes nutrients from the intestinal tract to other parts of the body. Of course, it is possible that Sterelny and Griffiths would want to say that digestion per se is not an adaptation, but is instead the interaction of various adaptations. However, this would be a contentious claim.

The reason why Sterelny and Griffiths think their problem holds for behavioral cases is that *where a psychological mechanism causes more than one behavior, changing one behavior will result in changes to the others* because the changes to one behavior have to be effected by changes to the mechanism that produces both behaviors. If a behavior B_i is produced by a mechanism M_{Bi}, and there is no one to one connection—i.e., M_{Bi} produces another behavior B_k—then if M_{Bi} changes such that B_i changes, then B_k will change also. Sterelny and Griffiths' claim is that if B_k also changes then B_i cannot count as mosaic.

To show how their claim might apply to a real sociobiological example, let us return to the case I mentioned in the introduction, of Hrdy's (2000) work on the evolution of conditional infanticide behavior. Hrdy claims that infanticide arose as a conditional strategy that human mothers could employ under extreme circumstances. Primate mothers rarely commit infanticide. Human infants require more energy and attention for longer childhoods that primate infants and human mothers can afford to raise very few of them. Making a "bad" choice regarding a human infant (in terms of birth spacing or health) is a more serious mistake for a mother from an evolutionary point of view than making the same mistake as regards a primate baby. Hence Hrdy thinks that infanticide is an adaptation to these new demands placed on human mothers; however, it only occurs under conditions of high stress, such as close birth spacing, famine, or infant deformity.

Sterelny and Griffiths' claim would be that conditional infanticide is unlikely to count as an adaptation in its own right. However infanticide occurs, it is supported by a variety of mechanisms that are involved in the production of other behaviors—mechanisms that are involved, say, in other aspects of reproductive behavior or more general-purpose practical reason. Change to those mechanisms to permit infanticide as such, say to the decision rules that permit general-purpose practical reason, would ramify to the other behaviors controlled by those decision rules. Non-infanticide behavior caused by those changed decision rules would perhaps no longer be appropriate or rational. Hence the evolutionary benefit of being able to choose infanticide would be masked by the damage caused by behaving ineffectively or inappropriately as regards other matters.

There are two main problems with this view, and thus with Premise 4 of the strong version of Sterelny and Griffiths' argument. The first is that it is not clear why having one mechanism subserving two behaviors means that changing one behavior necessitates changes in the other: i.e., why it is that

https://doi.org/10.1086/381410 Published online by Cambridge University Press

when M_{Bi} changes B_i that B_k must change as well. If it is not clear that M_{Bi} must change B_k if it changes B_i, then it is not clear that functional complexity *does* mean that behaviors are inevitably connected to each other. If they can evolve independently after all, Premise 4 falls through and the strong interpretation of Sterelny and Griffiths' argument fails.

The second objection is that Sterelny and Griffiths' argument here relies on too strong a reading of the quasi-independence criterion. Even if behaviors are connected via a complex psychology, Sterelny and Griffiths have not shown behaviors are sufficiently strongly connected that natural selection cannot in principle get to work on them.

2.4.1 Mechanisms and Connections Between Behaviors. I will address the first problem with Premise 4 first: why might we think that if M_{Bi} changes B_i it must change B_k? As things stand, we do not know much about how psychological mechanisms work in humans or animals to any level of precision. What we do know is that, as Sterelny says, human cognition is functionalist in character—i.e., there are a number of mental mechanisms—be it multitudes of "modules" or a "central processor" with peripheral "transducers"—that take in informational inputs, process the information and output some kind of behavior. Sterelny and Griffiths claim that some of these mental mechanisms control multiple behaviors.

It seems Sterelny and Griffiths are right that the evolution of behavior is fairly straightforward where you have a one to one connection between mechanisms and behaviors. You change B_i by changing M_{Bi}. To use an artificial example—suppose that a species of bird has a cognitive mechanism that processes certain sorts of perceptual and other inputs, and outputs an alarm-calling behavior in response. Suppose the mechanism fairly rigidly controls the behavior. Then changes to the alarm call must be created by appropriate changes to the specific mechanism that controls it. Nothing else need change.

What would happen in the case where one mechanism causes many behaviors or where many mechanisms cause many behaviors is less clear. I assume that a single mechanism is able to control the production of different behaviors by being sensitive to different inputs; I assume that for each unique behavior there is a unique set of inputs to the relevant mechanism(s) that produces it. To change one behavior, you have to change the sensitivity of the relevant mechanism(s) to the relevant set of input(s). For example, M_{Bi} produces B_i in response to one set of inputs and B_k in response to a different set of inputs. M_{Bi} already responds differently to different inputs to create the two different behaviors. Why should we in principle expect variations in M_{Bi}'s production of B_i to produce variations in its response to B_k?

https://doi.org/10.1086/381410 Published online by Cambridge University Press

Take the following toy example. Suppose we have a mechanism that is implicated in responses to the emotions of others—say by raising or reducing the aggression level. The algorithm this mechanism supports raises the aggression level if another mechanism (say, a facial-expression recognizer) sends a signal that the face recognized is angry. It lowers the aggression level if another mechanism indicates that the face is sad. In essence there is a "branch" in the algorithm that allows the mechanism to be sensitive to two different inputs—sad-face and angry-face—and produce two different outputs. Suppose there is pressure to evolve the aggression-response so that the mechanism inhibits the aggression response to an angry face if the face is recognized as kin. The mechanism evolves to incorporate inputs from a kin-recognizer regarding whether the face is kin. The algorithm acquires a new instruction on its angry-face branch which tells it to reduce aggression if the face is a family member, and raise aggression if it is not. The signals from the kin-recognizer are redundant on the sad-face branch. This change would allow the mechanism to evolve sensitivity to family members without affecting the other kinds of aggressive behavior.

Or the separate changes to particular types of behavior might occur by other, more straightforward means such as changes to behavior-specific representations. Consider again our real sociobiological example, Hrdy's work on infanticide behavior. Hrdy is not at all specific about what she envisions the psychology of infanticide to be like. According to Hrdy, infanticide may occur by means of a combination of maternal ambivalence in the few days after birth (Hrdy 2000, 487–488), together with some sort of reasoning on the part of the mother regarding her circumstances and in light of her desires and goals. One way in which independent evolution of infanticide could be achieved is by women acquiring representations of infanticide-specific goals or desires (e.g., the desire not to have babies too close together) that are operated upon by general-purpose algorithms. These representations could evolve separately. Changes to infanticide specific representations would affect infanticide related behavior, and be unlikely to affect anything else.

To sum up, mechanisms that handle more than one behavior must be able to deal with different sets of stimuli by producing different outputs, and this might be achieved in a variety of different ways. The very fact that mechanisms have to be sensitive to different inputs suggests that there is a degree of independence in the means by which mechanisms handle those inputs. If those means are independent, then surely it will be possible to change the way a mechanism handles one behavior without automatically affecting the others. At the very least we can say that we do not know that changing M_{Bi} to change B_i *must* also change B_k.

https://doi.org/10.1086/381410 Published online by Cambridge University Press

2.4.2 The Quasi-Independence Criterion and Behavioral Connectedness. I now want to present the second objection to Premise 4. Even if Sterelny and Griffiths are right that where more than one behavior is produced by the same mechanism, changing one behavior means changing the other, they still have not made their argument work. Sterelny and Griffiths need to show that a one-to-one connection between behaviors and mechanisms is *necessary* for a behavior B_1 to meet Lewontin's quasi-independence criterion. In other words, Sterelny and Griffiths need to show that not only is there a connection between B_i and B_k via their mechanism M_{Bi}, but that the positive fitness benefit of the change to B_i is outweighed by the negative impact on B_k. The quasi-independence criterion shows that in order for a trait to count as quasi-independent there must be a way for the trait to evolve without causing *countervailing* changes elsewhere. The quasi-independence criterion does not simply say that changes to one trait T_i must not ramify to others $T_1 \ldots T_n$; it says that the changes to $T_1 \ldots T_n$ must not *countervail* the positive fitness effects of the change to T_i. Sterelny and Griffiths need to show that those changes are *inevitably* negative enough that they outweigh the fitness contribution of the positive changes, and that consequently the behavior cannot evolve. However that is not always the case; there are examples in the evolutionary literature of behaviors that are clearly genetically connected, but where one behavior is still able to evolve because the detrimental effect of one trait did not outweigh the positive contribution of the other.

One such example can be found in Stevan Arnold's (1981) work on prey identification in garter snakes (see also Brodie 1989, 1993). Newborn snakes from a coastal and an inland population in California were tested for their response to the scent of various kinds of prey they were likely to meet and hunt as adults: the inland population generally eats amphibians and fishes and the coastal population feeds on terrestrial slugs. They were also tested for their response to leeches and poisonous salamanders. There was a strong correlation between response to slugs and response to leeches, though the slug eaters showed significantly more response to both the slug and the leech than the frog and fish eaters. However, eating leeches is dangerous for garter snakes, since when swallowed, the leech stays alive in the gut of the snake. The strong correlation between behavior exhibited by naive newborns suggests pleiotropy—whatever gene(s) create the response to the slug also create a response to the leech, probably due to the chemical receptors for the scent of one species also being sensitive to the scent of the other species.

The slug response behavior has likely evolved in the coastal species because the snakes do not encounter many leeches; hence the selection against the correlated leech response is low enough not outweigh the positive benefit of detecting slugs. In the inland variety, which encounters

https://doi.org/10.1086/381410 Published online by Cambridge University Press

a lot of leeches, selection against response to leeches is strong enough to lead to a decrease in the connected response to slugs. In other words, slug response behavior had undergone natural selection in both populations of garter snake, despite a strong pleiotropic connection with the (dangerous) leech response behavior. To relate this back to our argument—in the coastal population the slug-response behavior evolved because the connection with the leech response did not lead to a countervailing negative fitness benefit—i.e., in this case the slug response met the quasi-independence criterion despite the fact that it is *genetically* connected to another trait that has a different effect; even though the slug response behavior has an *even stronger connection* to leech response behavior than just sharing an underlying mental mechanism, the slug response was able to evolve in the face of the leech response.

3. The Weak Form of Sterelny and Griffiths' Argument. It seems that the stronger form of Sterelny and Griffiths' argument faces some serious problems. However, I think one might interpret the argument in a weaker way. This form, I think, does go through; however, it does not really undermine the project of sociobiology in the way that Sterelny and Griffiths seem to be seeking. I take it that this form runs roughly as follows:

1. Sociobiology tries to discover behaviors that are adaptations.
2. If a behaviors is to be an adaptation it must be mosaic, i.e., it must meet Lewontin's (1978) quasi-independence criterion.
3. Some important human behaviors have connections that violate the quasi-independence criterion.
4. It is very hard to determine which behaviors have psychological connections that violate the quasi-independence criterion.
5. Therefore, it is very hard to determine which behaviors can be adaptations.
6. Therefore, it is very hard to do sociobiology.

If this is what Sterelny and Griffiths have in mind, then they are really presenting a caution along similar lines to Gould and Lewontin's (1978) criticism of Panglossian adaptationism—that hidden constraints prevent some behaviors from being adaptations. Sterelny and Griffiths argue that natural selection cannot always do what sociobiologists claim it can; some human behaviors have hidden connections that prevent them from evolving in just any direction. Many of these traits might be the kind of interesting behaviors which sociobiologists want to explain. If this is an accurate representation of their view then Sterelny and Griffiths have a point: behaviors may have hidden constraints upon their evolution in the same way that certain physical traits might have. Discovering those

https://doi.org/10.1086/381410 Published online by Cambridge University Press

connections can be very difficult, and hence it can be hard to determine, in any given case, whether we are looking at a real adaptation or not.

However, if this is how we should take Sterelny and Griffiths' argument, its capacity to support their larger conclusions is limited. First, the problem of determining the presence of psychological constraints does not prevent us from doing sociobiology. Second, weakening the argument this way does not make evolutionary psychology much better off than sociobiology; it falls afoul of the same problem (albeit in a weaker form).

3.1 The Weak Form Does not Undermine the Sociobiological Project. Sterelny and Griffiths are right in the concern expressed in my weaker interpretation of their argument. It is true that we cannot expect any trait to be straightforwardly independent of the other parts of the organism in which it occurs; we cannot expect it to be able to evolve in just any direction; and we cannot expect changes in one trait to leave the rest of the organism unaffected.

However, saying that behaviors *may* have hidden constraints upon them is a far cry from saying that behaviors inevitably have these constraints, and especially from saying that these constraints render most behaviors incapable of having evolutionary histories of their own. Sociobiologists may need to be careful to watch for these hidden constraints when describing behavioral adaptations, but the existence of such constraints does not make it impossible to pursue sociobiology. Sociobiologists have tools that permit the identification of adaptations; clearly they need to make sure that these tools are capable of distinguishing connected behaviors that cannot be adaptations from mosaic behaviors that can.[7] But there is no reason in principle why this should not be possible.

3.2 Evolutionary Psychology Faces Similar Problems. The second problem with the weak interpretation of Sterelny and Griffiths' argument is that it does not help their argument for replacing sociobiology with a form of evolutionary psychology.[8] If the argument can be made in the case of sociobiology, a similar argument can be made in the case of evolutionary psychology. The reason is that mental mechanisms are also produced by a substrate—the neurology of the brain. The brain's physical

7. An example of such a tool would be the optimality model. For excellent discussions of how sociobiologists use optimality models to incorporate constraints of other kinds, see Maynard Smith (1978) and Parker and Maynard Smith (1990).

8. Sterelny and Griffiths are clear that they do not wish to replace sociobiology with evolutionary psychology as it currently is practiced, since that form of the project has commitments with which they do not agree—notably the idea that the mind is composed largely of mental 'modules'. What they want is a science that investigates the origins of human psychological mechanisms.

https://doi.org/10.1086/381410 Published online by Cambridge University Press

structures and processes are responsible for supporting algorithms that process information and produce behavior. In order to change some characteristics of a mental mechanism, odds are you have to change something about the way the brain is structured. For all we know there may be constraints on how one can change neurology to change mechanisms: for example, mechanisms might share neurological processing space—changing the way one operates might require that that processing space change its character and thus also affect the performance of another mechanism as a consequence. It is certainly the case that many parts of the brain seem to have a role in multiple (if related) cognitive capacities.[9] It is possible that changing the genes that produce these brain structures would similarly result in changes to multiple cognitive mechanisms.

The truth is that we do not know enough about how mental mechanisms get instantiated in neurology, nor how neurological changes will create psychology—perhaps mental mechanisms are plastic enough that genetic changes resulting in neurological change can preserve one cognitive mechanism and not another. But the point is clear enough. *As far as we know* there is reason to think that mental mechanisms *might not* be mosaic, in the same way as there is to think that behaviors might not be mosaic. Because mental mechanisms are instantiated in a neurology that also might make them "connected" traits, i.e., changing one mechanism may have a countervailing deleterious effect on one or more of the others. If the potential connectedness of behaviors raises a serious impediment for sociobiology, then the potential connectedness of mental mechanisms ought to raise a serious impediment for evolutionary psychology.

4. Can Behaviors Be Adaptations? Clearly then, Sterelny and Griffiths argument that behaviors cannot be adaptations falls through. Some of the behaviors that are interesting to sociobiologists could have evolved by natural selection; or at least, we need not expect prima facie there to be psychological constraints preventing all interesting behaviors from evolving. In this section I wish to go a step further and present some reasons for thinking that both behaviors and mental mechanisms can be the targets of natural selection; i.e., that that it is possible for humans to possess separate behavioral and cognitive adaptations.

The reason why I think I need to emphasize this point is that the distinction between behavioral and cognitive adaptations is occasionally confused. Sometimes, even in the literature on sociobiology and evolutionary psychology, attributions of behavioral and psychological

9. One example might be the amygdala—this seems to have some role in producing a response to sadness and fear in others (Blair et al. 1999) and in controlling mood and anxiety in the subject (Davis 1992).

https://doi.org/10.1086/381410 Published online by Cambridge University Press

adaptations are made interchangeably. For example, in Martin Daly and Margot Wilson's (1996) discussion of the evolution of homicide the homicide adaptation is sometimes described as if it were a behavior and sometimes as if it were a mechanism. Mechanisms and behaviors seem to be regarded as a single 'package', not as components that can evolve independently of each other.[10]

The primary reason for holding this view, I imagine, is a perception that the fitness effects of mental mechanisms and the behaviors they generate are equivalent. Obviously, the central evolutionary function of a mental mechanism is to produce certain kinds of behavior; and that behavior is the main means by which the mechanism interacts with its environment. It ought to be easy to argue, then, that mechanisms evolve because they produce fit behaviors; they are fit because they make fit behaviors—and thus it seems that the fitness effects of any mechanism are effectively equivalent to the fitness effect of the behaviors it produces. This suggests that the mechanisms and their behaviors are a unit from an evolutionary point of view—their contribution to the fitness of an organism is identical. I am going to give some reasons for thinking that this view is incorrect.

4.1 Mechanisms Can Be Adaptations Independent of Behaviors. One reason for thinking that both behaviors and mechanisms can be adaptations in their own right is that, contrary to the objection that I outlined above, the contribution to fitness of cognitive mechanisms (at least, in ancestral environments) is *not* equivalent to the contribution to fitness of the behaviors they produce in those environments. It is true that mental mechanisms in some sense produce behaviors, but their fitness value is not equivalent to the fitness value of any particular behavior they produce.

There are two main reasons to think this. First, many mental mechanisms contribute to more than one behavior, so their fitness is not equivalent to the fitness of any given behavior. Second, the fitness value of any mechanism is still not equivalent to the overall fitness of all the behaviors it produces because that mechanism will have characteristics of its *own* that make it fit independent of its production of useful behaviors: for example, processing efficiency, space efficiency, energy efficiency, and so forth. A mechanism possessed by some organism O might evolve in its energy efficiency, for example, without there being any significant evolution in the behaviors it produces—the energy changes increase O's fitness because they free up the energy reserves for other uses inside of O's body. Improvements in space efficiency might free up neurological

10. This view may, indeed, explain why evolutionary psychology is sometimes viewed as little more than sociobiology with a new name—their subject matter is regarded as the same in some sense.

https://doi.org/10.1086/381410 Published online by Cambridge University Press

tissue for the instantiation of new, useful mental mechanisms; or it might simplify the developmental process of that mechanism, making it less likely that the process could be disrupted; and so forth.

4.2 Behaviors Can Be Adaptations Independent of Mechanisms. Similarly, I think it makes sense to speak of behaviors evolving independently of mental mechanisms. This is because any behavior B_i is capable of being produced in different individual organisms by a variety of different mental mechanisms or sets of mental mechanisms.[11] Just as the possession of striped fur or a digestive enzyme might be generated by different genes in different organisms, so a behavior might be produced by different kinds of mechanisms in different organisms. For example, suppose in a population of people there exist two different xenophobia mechanisms that make their possessors xenophobic to those who speak differently from them. Both mechanisms use approximately the same decision rules (e.g., become aggressive when someone's speech pattern does not match the local dialect representation) and representations (e.g., a representation of the structure of local speech learnt in childhood), and hence both produce almost identical behavior under the same conditions. Both make people aggressive when presented with others who speak a non-local dialect. However, one mechanism (X_1) is very space efficient and makes fewer demands on its possessor's metabolism. The other (X_2) is instantiated untidily in the brains of those that possess it—it uses up lots of space, is demanding in its use of energy, and so forth. In that case we have the same behavior produced by two clearly different mechanisms.

I am claiming that where the same behavior is produced by multiple mechanisms, we can have cases where a behavior evolved by natural selection, whilst the mechanism that supported it had evolved by some other means. Suppose we have a small population of organisms who exhibit a variety of behaviors $B_1 \ldots B_n$ given the same set of stimuli. One of these behaviors B_i is clearly more effective at increasing the survival and reproduction of those organisms which possess it than the rest of $B_1 \ldots B_n$. B_i, however, can be reliably produced without fitness-affecting variation by a variety of mechanisms or mechanism groups $M_1 \ldots M_n$. Suppose that for reasons independent of the production of B_i, some mechanism M_i of $M_1 \ldots M_n$ is more fit than the rest of $M_1 \ldots M_n$. Suppose the selective advantage of exhibiting B_i relative to $B_1 \ldots B_n$ is greater than the selective advantage of possessing M_i relative to $M_1 \ldots M_n$. Then B_i could evolve to

11. We can think of the same B_i being produced by different mental mechanisms as long as B_i is taken to be the disposition to behave in a particular way (i.e., to produce certain specific bodily movements) given certain specific conditions.

https://doi.org/10.1086/381410 Published online by Cambridge University Press

fixation quickly—but underlain by a polymorphism of $M_1 \ldots M_n$. Then some mechanism M_k could evolve by genetic drift in the small population despite the smaller selective advantage of M_i. Under these conditions it would make sense to say that B_i was an adaptation but M_k was not.

Let's apply this to our earlier example. Imagine that the population in which X_1 and X_2 exist is very small. Those people that possess X_1 and X_2 are xenophobic to strangers, and in their environment this is beneficial from an evolutionary point of view; it makes them suspicious of people who might be enemies, or competitors for resources and so forth. In that case xenophobia evolves by natural selection—the whole population ends up behaving aggressively to those who don't speak the local dialect. However, this also means that every member of the population now possesses either X_1 or X_2. Now X_1 is more efficient than X_2; however, because the fitness difference is not very large, and because the population is small, X_2 evolves to fixation by genetic drift rather than by natural selection. It looks like xenophobia is an adaptation for the population in question, but X_2 is not an adaptation.

5. Conclusion: The Project We Need. Having given some reasons to think that both mechanisms and behaviors are adaptations, I now want to consider the implications of this for our choice between sociobiology and evolutionary psychology. It seems obvious that if both mechanisms and behaviors can be adaptations, then there is nothing inherently wrong with the idea of either a kind of sociobiology or an evolutionary psychology. However, I think that the way that these two projects are being currently carried out—as, in essence competitor projects (Cosmides and Tooby 1987)—may not be maximally helpful to the ideal of uncovering the evolutionary histories of human behavior and mental mechanisms. In order to decide how natural selection has acted in the evolution of human nature, we need to know how it has been acting on both mechanisms and behaviors. For this purely practical reason, it might be better to integrate the two projects.

Another reason for integrating the two projects is the usefulness of information about psychology to the investigation of behavior and vice versa. One way in which knowledge about psychology could assist with determining questions of behavioral evolution is in solving problems of atomization: determining which "packages" of behavior are the individual targets of selection and which are not. Behavioral "packages" could, for example, be identified by seeing which discreet groups of mental mechanisms produce which sets of movements an animal makes. Information about the structure of the underlying psychology that produces a behavior would also be useful for determining cases where heavy inter-mechanism connections have prevented the evolution of one behavior independently

https://doi.org/10.1086/381410 Published online by Cambridge University Press

of others. Knowledge about how mental mechanisms work would allow us to determine if changes to one behavior might make highly evolutionarily significant changes to another behavior. In the same way, knowledge of behavior can inform questions about the evolution of psychology and help to determine the existence and function of psychological mechanisms. Discontinuities in behavioral capacities suggest the computational and functional limits of psychological mechanisms or sets of psychological mechanisms.

Despite the fact that mechanisms and behaviors are sometimes sufficiently evolutionarily independent of each other to count as independent adaptations, there is no question that they are connected. You cannot investigate mechanisms to the exclusion of behavior: the evolutionary function of a mechanism is ultimately to produce behaviors of certain kinds, and the methodology of evolutionary psychology requires a scientist to examine the *behavior* of human beings to determine the possible existence of certain mechanisms. Similarly, you cannot investigate the evolutionary origins of behaviors without reference to mechanisms. Human behaviors are produced by mechanisms; what behaviors can evolve depends greatly on the sort of mechanisms that are available and how the evolutionary process can alter them. Nevertheless, it seems clear that a science that integrates information about behavioral and psychological adaptations will find it easier to deal with these problems. I suggest that what we need is a ''Psychological sociobiology''—a science of human nature that integrates information about the structure of human psychology and the nature of human behavior in generating Darwinian histories for the elements of ''human nature.''

REFERENCES

Arnold, Stevan J. (1981), ''Behavioral Variation in Natural Populations. I. Phenotypic, Genetic and Environmental Correlations between Chemoreceptive Responses to Prey in the Garter Snake, *Thamnophis elegans*'', *Evolution* 35(3): 489–509.

Blair, R. J. R., J. S. Morris, C. D. Frith, D. I. Perrett, and R. J. Dolan (1999), ''Dissociable Neural Responses to Facial Expressions of Sadness and Anger'', *Brain* 122: 883–893.

Brodie, Edmund D., III (1989), ''Genetic Correlations between Morphology and Antipredator Behavior in Natural Populations of the Garter Snake *Thamnophis ordinoides*'', *Nature* 342: 542–543.

———(1993), ''Homogeneity of the Genetic Variance-Covariance Matrix for Antipredator Traits in Two Natural Populations of the Garter Snake *Thamnophis ordinoides*'', *Evolution* 47(3): 844–854.

Cosmides, Leda, and John Tooby (1987), ''From Evolution to Behavior: Evolutionary Psychology as the Missing Link'', in John Dupre (ed.), *The Latest on the Best*. Cambridge, MA: MIT press, 277–306.

———(1992), ''Cognitive Adaptations for Social Exchange'', in Jerome H. Barkow, Leda Cosmides, and John Tooby (eds.), *The Adapted Mind*. Oxford: Oxford University Press, 163–288.

Daly, Martin, and Margot Wilson (1996), ''The Evolutionary Psychology of Homicide'', *Demos* Dec. 8, 1996, 39–45.

https://doi.org/10.1086/381410 Published online by Cambridge University Press

Davis, Michael (1992), ''The Role of the Amygdala in Fear and Anxiety'', *Annual Review of Neuroscience* 15: 353–375.
Gould, Stephen J. (2000), ''More Things in Heaven and Earth'', in Hilary Rose and Stephen Rose (eds.), *Alas, Poor Darwin.* New York: Harmony Press, 101–126.
Gould, Stephen J., and Richard Lewontin (1978), ''The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme'', in Elliot Sober (ed.), *Conceptual Issues in Evolutionary Biology.* Cambridge, MA: MIT Press, 73–90. First published in *Proceedings of the Royal Society of London* B 205: 581–598.
Hrdy, Sarah Blaffer (1999), *Mother Nature: A History of Mothers, Infants and Natural Selection,* New York: Pantheon/Random House.
Kitcher, Philip (1985), *Vaulting Ambition*. Cambridge, MA: MIT Press.
Lewontin, Richard C. (1978), ''Adaptation'', *Scientific American* 239(3): 212–230.
Maynard Smith, John (1978), ''Optimization Theory in Evolution'', *Annual Review of Ecological Systems* 9: 31–56.
Parker, Geoffrey A., and John Maynard Smith (1990), ''Optimality Theory in Evolutionary Biology'', *Nature* 348: 27–33.
Sterelny, Kim (1992), ''Evolutionary Explanations of Human Behavior'', *Australian Journal of Philosophy,* 70(2): 156–172.
Sterelny, Kim, and Paul Griffiths (1999), *Sex and Death*. Chicago: University of Chicago Press.
Wilson, Edward O. (1975), *Sociobiology.* Cambridge, MA: Belknap/Harvard University Press.

https://doi.org/10.1086/381410 Published online by Cambridge University Press