PSYCHOLOGICAL ALTRUISM VS. BIOLOGICAL ALTRUISM: NARROWING THE GAP WITH THE BALDWIN EFFECT

Mahesh Ananth

Department of History, Philosophy and Political Science, Chicago State University, Chicago, Illinois, USA.

Mailing Address: 9501 S. King Drive, Chicago, Illinois, 60628. Phone: 1-773-995-2414. Email: mananth@csu.edu

Received 8 June 2004, revised 12 January 2005, accepted 20 June 2005

ABSTRACT

This paper defends the position that the supposed gap between biological altruism and psychological altruism is not nearly as wide as some scholars (e.g., Elliott Sober) insist. Crucial to this defense is the use of James Mark Baldwin's concepts of "organic selection" and "social heredity" to assist in revealing that the gap between biological and psychological altruism is more of a small lacuna. Specifically, this paper argues that ontogenetic behavioral adjustments, which are crucial to individual survival and reproduction, are also crucial to species survival. In particular, it is argued that human psychological altruism is produced and maintained by various sorts of mimicry and self-reflection in the aid of both individual and species survival. The upshot of this analysis is that it is possible to offer an account of psychological altruism that is closely tethered to biological altruism without reducing entirely the former to the latter.

1. INTRODUCTION

How do non-human organisms develop the traits that they do? Given that no biological trait or behavioral capacity develops independently of environmental factors and that development involves complex interactions among genes and between genes and environments, the question posed has proven to be a notoriously complicated one.[1] Consider further the issue of whether it is possible to understand how the human organism develops its complex traits and behaviors given the environment/gene matrix. If the leading question concerning the details of non-human animal behaviors has been a difficult one for biologists and ethologists to answer sufficiently, one can imagine the amount of confusion when similar concerns shift to human behaviors and the corresponding psychological states. Historically, one such complex behavior in both non-human animals and humans that has been of great interest to scholars in many fields of inquiry is *altruism*.

Elliott Sober, in particular, has written a number of articles and books attempting to elucidate the problems that have plagued this discussion on altruism (Sober, 1988,

---

[1] Ananthakrishnan (1970: 241) makes this point clear with respect to ethology: "Field study of animals has attracted fewer students in our universities. It involves careful observation which is time consuming and there is much difficulty in dealing with organisms which show a behaviour; the latter obviously restricts the scope of studies and call for specialized approaches."

Acta Biotheoretica (2005) 53: 217–239

1993a: 199–216, 1993b, 1994; Sober and Wilson 1998: 17–54 and 296–327). His analysis in these works is two-fold: (1) He makes an effort to "disentangle" concepts such as *biological altruism* and *psychological altruism*, both of which he claims have been imprecisely used in many arguments; and (2) after explaining the distinctions between the above concepts, Sober discusses whether and how these concepts are related to each other. Although Sober does not rule out the possibility of an evolutionary explanation of psychological altruism, he does appear skeptical of this possibility for two reasons (Sober, 1993a: 210–11):

1. Human psychological altruism is about preference structures, while biological altruism is singled out by differences in reproductive success (fitness).
2. The human mind should not be considered "as a mere passive implementation of evolutionary imperatives."

This paper will explore, through Sober's work, some of the details of the evolutionary account of altruistic behavior known as *biological altruism* and the human conventional sense of 'altruism' known as *psychological altruism*.[2] I will proceed in the following manner. First, I will provide a brief history of the debate on altruism that will culminate with an account of the problem under consideration. Second, I will provide an analysis of Sober's distinction between *psychological altruism* and *biological altruism* and how he applies this distinction to the problem under consideration.[3] Third, I will show that Sober's distinction between *psychological altruism* and *biological altruism* is neither (1) necessarily as sharp as Sober considers nor (2) that this distinction necessarily vindicates the impossibility of an evolutionary explanation of *psychological altruism*. Finally, I will argue, employing a theory known as the "Baldwin Effect," that both the behavioral adjustments that the human organism makes in its lifetime and how such adjustments are both retained and transmitted to its descendants are part of what an evolutionary account of human psychological altruism should reflect. From such a perspective, I will argue that human psychological altruism is more closely tethered to human/non-human biological altruism than some have considered.[4]

An interesting implication of this line of reasoning is that Sober's account of psychological altruism will be couched within a general pluralistic framework of human other-regarding behavior. Importantly, when I claim that the "gap" between psychological and biological altruism can be narrowed, I am making two claims: (1) these concepts are not as conceptually distinct as Sober suggests and (2) that biological altruism is (in

---

[2] Note that 'biological altruism' refers to other-regarding behavior in both humans and other non-human animals, while 'psychological altruism' refers only to the other-regarding behavior of humans. See footnote 4.

[3] A note to the reader: Sober's definitions of both kinds of altruism are controversial. In this paper, I will not attempt directly to challenge them. Rather, I will take them as given, and focus my criticisms on the implications of his analysis. For a general philosophical discussion on the altruism debate, see Wilson and Dugatkin (1992).

[4] I have been grouping human biological altruism with non-human biological altruism. I will take it for granted that this is a reasonable move, given our shared evolutionary history with other non-human species. This sort of biological altruism would be related to the basic shared biology associated with the basic capacity for shared emotional commitment. This topic of the evolution of commitment is discussed in Nesse (2001).

part) causally related to psychological altruism (and potentially vice-versa, if the Baldwin effect is thought to be plausible)—that is, natural selection for biological altruism generates psychological altruism as a mechanism. The result will be a possible narrowing of the gap between human psychological altruism and biological altruism.

2. ALTRUISM: A BRIEF HISTORY AND THE PROBLEM

When a bee sacrifices its life for the sake of the hive of which it is a part, there is a sense in which its sacrifice is thought of as "altruistic." That is, examples of altruism reveal behaviors that cost the donor time, effort, life, or other resources for the sake of some other recipient. This type of other-regarding behavior is quite common in the natural world. The sacrifices of social insects, thump warnings of beavers, sharing of blood amongst vampire bats, primate grooming rituals, and human other-regarding behavior for both immediate family and strangers are but a few examples (Griffiths and Sterelny, 1999: Ch. 8; Ridley, 1996: Ch. 12).

Yet, from the traditional evolutionary perspective, altruism is a difficult pill to swallow, because behaving altruistically lowers the altruist's fitness relative to those who are recipients of such behavior. Indeed, on the traditional view, natural selection should disfavor or exploit altruism for the simple reason that selfish individuals will pass on their "selfish genes" to subsequent generations. After enough life cycles have passed, very few altruists (if any) will remain. The question, then, is why would any individual organism—human or otherwise—behave in such a way as to incur serious costs to itself while benefiting some other creature(s)? Darwin was acutely aware of the problem and addressed it directly: "he who was ready to sacrifice his life, as many a savage has been, rather than betray his comrades, would leave no offspring to inherit his noble nature" (Darwin, 1882: 163).[5]

Still, 'human psychological altruism' frequently refers to the bestowal of a benefit by one person (a donor) to another (a recipient) in a selfless way. That is, when psychological altruists express concern for the welfare of others, they do so without any expectation of reciprocity (blood donation is used by Dawkins, 1989: 230). Given that humans are also the product of evolutionary processes, however, it seems reasonable to suggest that the more conventional idea of human psychological altruism ought to be understood in terms of biological altruism.[6] Yet, to bring human psychological altruism within the fold of biological altruism would be to ignore the conventional sense of 'psychological altruism'.

---

[5] Michael Ruse (1998: 445) succinctly makes the point as follows: "If one has two organisms, one putting effort into the welfare of others and the other devoting its labours exclusively to its own end (although it may accept the aid of others), then, all other things being equal, the altruism is doomed to rapid evolutionary extinction. Selection is a short-term process, in that it does not think of the long-range gains. Even though the group may benefit in the long-run from the altruist's behavior, selection will never let genes for self-sacrifice get established."

[6] This could be by way of kin-selection (Hamilton, 1964, 1975; Wade, 1980), reciprocal altruism (Trivers, 1971; Skyrms, 1996; Maynard Smith; 1964, 1982), or group selection (Darwin, 1882; Wynne-Edwards, 1962; Sober 1993a, 1994; Sober and Wilson, 1998; Queller, 1992).

Is it possible to include the conventional sense of 'human psychological altruism' within the framework of human/non-human biological altruism? To be sure, this question has generated a great deal of controversy and confusion dating as far back as Darwin's own reflections on the subject (Darwin, 1882: 82–93; Bradie, 1994: 15–55). Primarily, the issue of the relationship between human psychological altruism and human/non-human biological altruism is located in a more general concern about what degree the fundamental genetic principles of evolutionary adaptation can be applied to the social behavior of humans. The question could be framed as follows: to what extent can results derived from the study of biological altruism in human and non-human animals be extended to an analysis of human psychological altruism? I suggest that the following argument may help to set up what has proven to be a turbulent debate:

P1 Either it is possible that the gap between human/non-human biological altruism and psychological altruism is narrow or the gap between human/non-human biological altruism and psychological altruism is wide.
P2 If the gap between human/non-human biological altruism and psychological altruism is wide, then it must be the case that psychological altruism (unlike biological altruism) is the product of causes peripheral to evolution by natural selection.
P3 It is not necessarily the case that psychological altruism is the product of causes tangential to evolution by natural selection.
P4 It follows that it is not necessarily the case that the gap between human/non-human biological altruism and psychological altruism is wide.
C1 Therefore, it is possible that the gap between human/non-human biological altruism and psychological altruism is narrow.

Assuming that P1 and P2 are not themselves problematic in the above argument, I submit that the heart of the altruism debate has been understood in terms of explaining the relationship between the other-regarding behavior of human/non-human animals and the other-regarding behavior and concomitant psychology of humans in P3 and whether this explanation will support the conclusion. Restated, I want to know whether the biological explanation of human/non-human animal altruistic behavior can be extrapolated onto human psychological altruistic behavior so as to determine the soundness of the above argument. Some (Ruse and Wilson, 1986; Wilson, 1975; Richards, 1987) think that the explanation of human/non-human biological altruism can be extrapolated onto human psychological altruism; moreover, they consider the gap between human/non-human biological and human psychological altruism to be negligible, given the truth of P3. Others (Sober, 1993a; Kitcher, 1985; Nagel, 1978: 198–205), more conservatively, insist that it is not possible for such an extrapolation to be achieved. They argue that the gap is substantial, because P3 is false. In this essay, Elliott Sober will stand as the exemplar of this camp.[7] I hope to show that a middle ground between these two extreme positions is defensible.

---

[7] Of course, the accounts of the other members in this camp will differ in the details, but the conclusions they reach are similar. Notably, Sober and Wilson (1998: 296–337) *seem* to align themselves much more closely to the pluralistic picture I argue for in this paper, although they do not take any definitive stance. Most of this text is designed to refute the plausibility of psychological egoism.

3. SOBER'S DISTINCTION BETWEEN PSYCHOLOGICAL ALTRUISM AND BIOLOGICAL ALTRUISM

Sober distinguishes between *vernacular altruism* (hereafter referred to as psychological altruism) and *evolutionary altruism* (biological altruism) as a means of arguing that there is a distinct gap between the two concepts. According to Sober (Sober, 1988: 76–78), there are four elementary features of psychological altruists:

(1) they have minds,
(2) they have no concern with regard to their own reproductive success,
(3) they make no comparative analysis between acts, and
(4) when they confer benefits upon others, they do so in a non-instrumental manner.

In evolutionary biology, on the other hand, altruism has a rather different meaning than its psychological counterpart that was just considered above. Biological altruism has been understood as a function of benefits conferred by some individuals upon others that simultaneously results in the reduction of the fitness and reproductive success of those imparting the benefits. Biological altruism, in contrast to psychological altruism, has the following characterization in Sober's analysis (1988):

(1) "reproductive benefits" is the only criterion,
(2) a mind is not (necessarily) a criterion,
(3) a cost/benefit analysis and comparison between acts are always part of "reproductive benefits."

The point Sober is trying to make clear here is two-fold: (1) Biological altruism is a comparative term, whereas psychological altruism is concerned with particular motives of individuals. That is, psychological altruism is not a comparative term (Sober, 1988: 77). Deciding whether an organism is a biological altruist requires assessing not only the benefits it confers on its conspecifics but also what its conspecifics yield in return. (2) Biological altruism, unlike psychological altruism, should not be thought of as synonymous with donation. Donation is simply the giving away or conferring of benefits. Without a comparison of benefits donated and benefits received, determining whether an organism is a biological altruist is not possible. In Sober's own words, "every [biological] altruist is a donor, but not every donor is a [biological] altruist" (Sober, 1988: 83).

Having made clear Sober's distinction between psychological altruism and biological altruism, it can be asked how the two concepts might be related. Restated, what can be said about the relationship between human psychological altruism and non-human biological altruism (P3)? Sober offers two interesting possibilities. The first possibility is that human psychological altruism is an evolutionary "spin-off." That is, human psychological altruism was not a trait governed by natural selection, but a property that was made manifest as a result of selection for other mental properties. This view would suggest that human psychological altruism has no direct genetic component and is not directly adaptive (P3 is false). Unfortunately, Sober does not explore the details of what such a spin-off theory would look like.

The second possibility Sober offers, in the spirit of Darwin's own suggestion, is that human psychological altruism is a trait that was under the control of evolutionary forces. This means, for Sober, that human psychological altruism is a trait that is adaptive, has a genetic component, and is the product of group selection (Sober, 1988: 95–96).[8] Sober proposes that *if* human psychological altruism was under the control of evolutionary forces, then a group selectionist account (Sober and Wilson 1994, 1998) would be necessary to give support to a connection between biological altruism and psychological altruism as he describes them:

> Suppose for the sake of argument that we have evolved by group selection and that many of our behaviors exist because they are beneficial for the group to which we belong. How might Mother Nature have wired us up to get us to behave in this way? That is, which proximal mechanisms could produce behaviors that have this sort of distal explanation? One possibility is that we should be psychological altruists. We might have preferences concerning the welfare of others and we might in various circumstances accord those preferences stronger weight than the preferences we have about our own welfare (Sober, 1993a: 211).

Yet, Sober does not take seriously the above account. He offers the following reason:

> There is another role the mind can play, however, one that I think sociobiologists have underestimated. A characteristic may evolve for a given evolutionary reason, but then have consequences that would not be at all foreseeable if only the evolutionary explanation was taken into account (Sober, 1993a: 203).

Sober's misgivings reflect the fact that his account of psychological altruism requires that the truth of P3 remains sufficiently dubious in order to avoid the pitfalls of many of the reductionist/adaptationist sociobiological accounts. No doubt, Sober's concerns above should be taken seriously.[9] It is easy to see how the strategy of relying on evolutionary considerations could be abused. For example, there may be no obvious distal explanation of how humans have developed the ability to play music, write poetry, or create complex mathematical or logical systems. These characteristics, which are unique to the human organism, may very well be unexpected by-products of a highly evolved brain. Still, the possibility of defending the origin of human altruism in the way that Sober proposes in his second possibility is worth exploring. Let us turn directly to the "Baldwin Effect" for such a plausible explanation.

---

[8] Sober and Wilson's (1994 and 1998) theory of group selection and its relationship to altruism (human or non-human) is controversial and the topic of much discussion within the philosophy of biology. The details of this theory cannot be addressed here. Suffice it to say, Sober appears to be in favor of the first possibility where he views human altruism more along the lines of an evolutionary "spin-off." For an attempt to settle the score as to whether or not Sober and Wilson's theory of group selection is really "group selection," see Okasha (2001). Also, see Williams' (1966: Ch. 4) classic discussion on the possibility of group selection.

[9] This is not to suggest that it is commonplace amongst scientists (e.g., behavioral ecologists) to treat every human feature as an adaptation—so-called "just-so stories." The point is simply that one should be careful about what human features can reasonably be included under the umbrella of biological evolution. For those who are skeptical about such abuses, see Kitcher's (1985) careful analysis of many who have tried to extend evolutionary theory too far with respect to human behavior.

4. THE BALDWIN EFFECT

In 1896, the psychologist/developmental biologist James Mark Baldwin proposed a theory called "Organic Selection" as a needed integral element of Darwin's principles of variation and selection (Baldwin, 1896, 1980). In the past decade, there has been a concerted effort on the part of some scholars to revive some of Baldwin's insights, which they consider provide an accurate account of the tempo and nature of evolutionary change. These insights have been loosely brought together under the title of "The Baldwin Effect." This section will provide the elements of the Baldwin Effect by way of Baldwin's own thoughts and some of the interpretations that have emerged out of the contemporary Baldwinian resurrection. The subsequent section will then use the Baldwin Effect to make sense of human psychological and biological altruism.

To start, Baldwin was grappling with how to make sense of the modifications an organism undergoes in its lifetime (ontogenetic variations) and what, if any, relationship exists between such modifications and the historical development of the species (phylogeny) to which a particular organism belongs. With respect to humans, ontogenetic variations would include both mental phenomena and behavior, which according to Baldwin should be part of an accurate understanding of human evolutionary change.[10] Indeed, for Baldwin, variations in learned behavior are the catalysts for both developmental and evolutionary change (Depew, 2003).[11] The following passage from Baldwin

---

[10] It is worth noting that Darwin (1996: 97) also gave importance to behavior as a catalyst to evolutionary change. He tells us: "It must be admitted that habits whether congenital or acquired by practice "sometimes" often become inherited: instincts, influence, equally with structure, the preservation of animals; therefore selection must, with changing conditions tend to modify the inherited habits of animals. If this be admitted it will be found *possible* that many of the strangest instincts may be thus acquired." It would not at all be surprising if Baldwin was influenced by this passage, although there is not direct evidence to suggest that he was.

[11] The historical context of this discussion is that Baldwin was challenging the supporters of Lamarckian inheritance of "acquired characteristics" (e.g., Herbert Spencer and George Romanes). Lamarck and many of his supporters thought that changes in the external environment would cause changes in the internal structure of organisms. Baldwin offered just the opposite view. He put forth that an organism's congenital variations, to the extent that they influence what that organism does in its behavior, physiology, and morphology, are at the same time changing the organism's environment and the organism itself in the short-run, which in turn can modify the make-up of the species to which the organism belongs in the long-run through natural selection. Also, Baldwin was responding to preformationists who held on to the view that the seed or germ line contains a miniature replica of the parent (e.g., H. Boerhaave, J. Swammerdam, and G. Cuvier). A brief discussion of preformationism is available in Depew and Weber (1995: 41–42). Lastly, it should be pointed out that C. L. Morgan (1896: 305) shared a similar view to that of Baldwin. With respect to the role of learning and evolution, he vaguely states, "variation does seem in some cases to have followed the lines of adaptive modification, so as to suggest some sort of connection between them." See also Simpson (1953) and Galef (1987) for critical discussions of this history.

reveals the core of what is now being called the Baldwin Effect:

> The variations which we find available for physical inheritance are congenital changes; the utility of individual modifications is confined to their influence in screening, supplementing and preserving the natural equipment of individuals and species, and thus directing the course of evolution. We have no reason to depart from this position in the matter of mental variations and the education of the individual. Mental characters already congenital are inherited; and the plasticity, which intelligence carries with it, is a congenital character. There is no evidence of the transmission of the results of mental education and experience; but both physical and mental endowments and the variations arising in them are subject to continuous physical transmission. So far the consistent application of Darwinian principles. (Baldwin, 1980: 27–28).

Baldwin notes that ontogenetic changes aid in maintaining and modifying genetic endowments and producing potentially useful variation on which selection can work. In the same way, he thinks that mental variations and education, which are part of human ontogenetic development, also assist in influencing genetic endowment and variation. As Downes notes (2003: 35–36), “His [Baldwin’s] intuition was that what made an organism successful was the whole constellation of its physiological and behavioral traits, whether the behavior was acquired or instinctual.” In contemporary parlance, Baldwin thought that the unit of selection was the individual organism—both its congenital traits and the non-congenital features acquired during development.

The immediate question that comes to mind is how can the modifications just now described be secured in an organism’s lifetime? There are two parts to the answer of this question. First, is the idea that Baldwin labeled the “law of use and disuse” (Baldwin, 1896: 444).[12] ‘Use’ refers to movements and behaviors that prove favorable to an organism and are thus continually repeated and eventually retained. ‘Disuse’ refers to those movements and behaviors that are harmful to an organism and are eliminated by a lack of reinforcement.

The second means of securing ontogenetic modifications, according to Baldwin, is through imitation. Baldwin defines an imitative reaction as, “one which normally repeats its own stimulus” (Baldwin, 1894: 48). An example will help flesh out the details of this rather broad definition. Let us consider one species of Darwin’s famous finches, the large ground finch that inhabits the Galápagos Islands. Let us further suppose that El Niño is in full swing in the Galápagos Islands and, as a result, resources are scarce. Although most of the large ground finch’s food resources have been destroyed by the turbulent weather, there are a few kinds of nuts still available for consumption. Unfortunately, the shells on the remaining nuts are very hard. The large ground finch has neither the beak strength nor the force to crack open the nut in order to get at its meat. Suppose the finch randomly attempts a multitude of strategies to crack open the nut, but all attempts fail. Finally, after much effort the ground finch happens to push a nut near a rock, which it

---

[12] This is not an exhaustive account of all the factors Baldwin thought were involved in ontogenetic development. For example, Baldwin also categorizes the different ways that ontogenetic modifications can occur (See Baldwin, 1896: 443–444). Of these factors, Baldwin considers imitation to be the most important in association with the “law of use and disuse.” Thus, special attention will be given to this factor of imitation.

then uses as a "plate" on which to crack open the nut (as opposed to the softer ground). The result is that the nut cracks open.[13]

Keeping in mind the "law of use and disuse," Baldwin would affirm that for the development (or survival) of the ground finch in our example it is crucial that useful reactions, such as pushing the nut toward the rock, be repeated in order to "outweigh the reactions which are damaging and useless" (Baldwin, 1894: 29). If reactions similar to those performed by the ground finch are immediately duplicated, these reactions will be retained (due to their usefulness) and will sustain the development of the bird. Thus, in general, we should expect that natural selection will favor genotypes that allow organisms the capacity to react adaptively with their surroundings and disfavor those organisms with less adaptability. Baldwin calls these kinds of repeated reactions imitative reactions.[14]

Baldwin concludes that the purpose of imitation is to fight off difficulties that arise in the short-run and keep the organism alive so that natural selection can come along in the long-run and turn those (or other related) ontogenetic adaptations into congenital instincts. As Depew summarizes (2003: 8), "through the action of natural selection working on the whole organisms, ontogenetic adaptations will eventually be fixed by germ-line shifts that reduce the contingencies to which presumably adaptive learned behavior is exposed." Baldwin stresses just this kind of active participation of the organism as it attempts to overcome local environmental stresses and how this affects the species, respectively:

> The imitative function, by using muscular co-ordination, supplements them, secures adaptations, keeps the creature alive, prevents the 'incidence of natural selection', and so gives the species all the time necessary to get the variations

---

[13] I have recently discovered that David Papineau (2004: 11–12) uses a similar example by way of the herring-gull practice of opening shellfish by dropping them on hard surfaces.

[14] In Baldwin's system, there are three kinds of imitative reactions, ranging from most rudimentary to highly complex. All three kinds are designed to "prevent the incidence of natural selection" and reflect a complexity in consciousness. They are categorized as: (1) simple contractility, (2) psychological or cortical imitations, and (3) plastic or secondarily subcortical imitations (Baldwin, 1894: 48–50). The first kind of imitative reaction is simply the contingent or chance variations that are copied due to their usefulness (e.g., the ground finch). This kind of imitation does not require any kind of complex intellectual stimulation for its use and retention. The second kind of imitative reaction, psychological imitation, is the copying of actions or words of another upon observation. Such imitations require a rudimentary kind of consciousness. The third class of imitative reactions falls under the title of "secondarily-subcortical imitation." These imitations include the two previous kinds of stimulus-repeating reactions except that they have been transformed into habit. That is, actions that were once either the product of accidental copying or actions or words that were the result of low-level conscious copying are retained through habit. Moreover, the cortical imitations are no longer in direct consciousness due to habituation. Included in this category are more complex social phenomena such as "the imitation of facial and emotional expression, moral influence, sympathy, and personal rapport" (Baldwin, 1894: 50). Note that some of these Baldwinian ideas about imitation (and the evolution of culture itself) may be vindicated by current research on the evolution of the mirror neuron systems in humans (See Ramachandran, 2003: 126–27).

> required for the full instinctive performance of the function (Baldwin, 1896: 448).

Thus, (in association with the law of use and disuse) it is the ontogenetic modifications that are retained predominantly through one of the three forms of imitation (see footnote 15) that led Baldwin to give the following definition of 'Organic Selection':

> The accommodations and modifications of the individual serve as a supplement or screen to his endowment; and in the course of time the endowment factor, by variation simply, without resort to the actual inheritance of acquired characteristics, comes to its perfection. This result of the 'coincidence' of modification and variation in guiding the course of evolution has been called 'organic selection' (Baldwin, 1980: 18).

Organic selection, then, is Baldwin's theory of how organisms engage in short-term problem solving. That is, 'Organic selection' is the combination of congenital factors in association with newly secured characteristics (accommodated through use and disuse) that allow an organism to solve immediate and pressing problems in its environment. In contrast, those organisms that must rely on congenital factors that do not allow for much adaptability, quite frequently, do not survive. The further question, then, is how do these ontogenetic modifications relate to phylogeny?

Baldwin's answer as to how ontogenetic modifications of an organism (tokens) are connected to its species (types) is clearly stated:

> The variations which were utilized for ontogenetic adaptation in the earlier generation, being kept in existence, are utilized more widely in the subsequent generation. Congenital variations, on the one hand, are kept alive and made effective by their use for adaptations in the life of the individual; and, on the other hand, adaptations become congenital by further progress and refinement of variation in the same lines of function as those which their acquisition by the individual called into play (Baldwin, 1896: 447).

As a way of illustrating the Baldwin Effect, two examples—one contrived and one actual case—will be provided. First, suppose we have a population of deer. Assume that part of the defense mechanism for this species of deer is its ability to zig-zag while running at high speeds. Now, add an increase in predation in this environment (imagine that a new predator invades the environment). As a result of this increase in predation, a few members of the deer population exhibit the ability to observe their immediate vicinity while drinking water.[15] It turns out that such a behavioral variation is useful in terms of early detection of predators—an accidental behavioral improvement that was retained through both imitation and use. According to Baldwin, such variations represent design improvements if they prove useful and are retained by the organism by habit and imitation (Baldwin, 1894, 1896), and are slowly refined after many generations

---

[15] In modern parlance, these kinds of ontogenetic changes fall under the general label of "phenotype plasticity." A phenotype is the body plan an organism has as a result of the interaction between its genes and its environment. The more malleable an organism's phenotype is, the better the organism is able to accommodate and withstand the "storm and stress of the physical influences of the environment, and of the changes which occur in the environment" (Baldwin, 1896: 447).

to the point where such variations become congenital through natural selection [16] (or that such an ontogenetic modification might allow the congenital retention of some other characteristic, such as improved vision or hearing).[17]

Second, a genuine case that can be reasonably understood under a Baldwinian framework is the selection for lactose tolerance. In mammals, the ability to breakdown lactose (a complex sugar in mammalian milk), by way of the production of lactase enzymes, ceases soon after weaning. When adult mammalian digestive systems attempt to breakdown lactose (as a result of ingesting milk products), they exhibit various sorts of gastrointestinal distress (e.g., abdominal cramps, bloating, weight loss, malnutrition, etc.) Yet, many sub-populations of adults (e.g., Danes, Finns, Hungarians, Mongols, northern Indian groups, Tutsi & Fulani pastoralists), where dairying has a long history, have shown the ability to digest milk—that is their lactase enzymes continue to function beyond the weaning years. Indeed, there is a close relationship between the frequency of lactose intolerance in a population and whether or not the population engaged in intensive dairy

---

[16] Godfrey-Smith, (2003: 54) offers a similar summary of the Baldwin Effect: "Suppose a population encounters a new environmental condition, in which its old behavioral strategies are inappropriate. If some members of the population are plastic with respect their behavioral program, and can acquire in the course of their lifetime new behavioral skills that fit their new surroundings, these plastic individuals will survive and reproduce at the expense of the less flexible individuals. The population will then have the chance to produce mutations that cause organisms to exhibit the new optimal behavioral profile without the need for learning. Selection will favor these mutants, and in time the behavior which once had to be learned will be innate."

[17] One might be concerned about how a learned behavior could become congenital. One plausible answer is what C. H. Waddington called "genetic assimilation" (see Waddington, 1942, 1952, 1953a, 1953b, 1961). This answer comes in three parts. First, the idea is that a novel "range" of a norm of reaction is made manifest when environmental shifts reveal a previously unexposed portion of the reaction norm (the ability to look up while drinking water in the face of heavy predation), which is kept somewhat stable by means of learning. Second, there are mutation, inversion, or crossover events (in the long-run) that allow for the production of the learned behavior at the genetic level with corresponding selection for the mutations, inversions, or crossovers (see Downing, 2004). Third, stabilizing selection secures the new norm of reaction in the direction of the environmental change. So, with respect to the deer example, the initially learned behavior allows the deer to withstand the abnormal increase in predation in the short-run by way of the relevant norm of reaction—in conjunction with different behavioral mechanisms such as stimulus enhancement, goal emulation, and blind mimicry—that is made manifest (see Papineau, 2004: 5; Shettleworth, 1998: 425–474). Then, when the relevant mutation, inversion, or crossover event(s) occur, there is selection for those genetic events that make the-looking-up-while-drinking-water-to-avoid predators-behavior as a stable part of the genome. Alternatively, it could be another characteristic (e.g., improved hearing or sight), as a result of genetic assimilation, that selection favors. So long as there is a close enough "link" between the learned behavior(s) and the characteristic(s) that is selected in the long-run, the Baldwin effect is in play (see also Schlichting and Pigliucci, 1998; Feldman and Cavalli-Sforza, 1989).

farming. Low levels of lactose intolerance are found in European populations with a long history of dairy farming, and highest levels in populations of Asian ancestry who were not dairy farmers (see Cavalli-Sforza, *et al*., 1994; Harrison, *et al*., 1988; Relethford, 1997). So, as a result of dairy farming activities *qua* behaviors, which were retained by both imitation and use and disuse, there was selection in favor of those genes to continue in the production of the lactase enzyme—a phenomenon known as adaptive enzyme formation (Ehrman and Parsons, 1976: 50–51).[18] The point of the example is that a change in behavior (farming practices and dietary habits) resulted in a change in gene expression, which is able to secure reliably those features associated with the learned behaviors. As Baldwin suggested, behavioral changes help the organism in the short-run, while selection helps the species in the long-run.[19]

The implications of the above examples are as follows. One, the ability of an organism to solve problems in its lifetime changes the competitive environment for its offspring (Depew and Weber, 1995: 275–297).[20] Second, genotype alterations via natural selection can possibly follow the lead provided by successful ontogenetic modifications. Third,

---

[18] I have intentionally used this mammalian example, giving emphasis to humans. The physiological changes of the stickleback fish are another *possible* example of the Baldwin Effect. Specifically, the norm of reaction shifts (or continued regulatory gene activity), as a result of dietary changes, have been detected (in laboratory conditions) in stickleback fish. That is, there is a moderate shift away from the phenotypic traits of the Limnetic form to the phenotypic traits of the Benthic form (and vice-versa) when their respective diets are switched. This change in environment/feeding behavior produces a corresponding norm of reaction change and a concomitant phenotypic change. For the details of this example and related genetic details about the stickleback fish, see Day *et al*., 1994 and Colosimo *et al*., 2004. This case is complicated by the fact that the observed changes may only be norm of reaction changes, resulting in no change at the genetic level. If this is the case, then this might not be a Baldwin-type effect. Yet, with respect to the stickleback (and other examples), West-Eberhard (2003: 169) notes, “The independent evolution of regulation and form means that there can be separate genetic accommodation of regulation and form of novel phenotypes. Once a trait is widespread in a population, further rounds of induced change and accommodation may occur when new inputs cause *modifications* of the regulation and/or form of the trait.” It appears that such accommodation, as a result of behavioral modification, occurs in the stickleback case, suggesting that it is a genuine Baldwin effect.

[19] Now, some scholars have argued that the dairy farming example is not a genuine Baldwin effect (Godfrey-Smith, 2003; Papineau, 2004), while others (Dennett, 2003; Deacon, 1997) insist that it is. The issue here is that the former insist that the feature that gets genetically assimilated must be the very same feature that was initially learned, while the latter insist that the feature that gets genetically assimilated needs to be similar (not identical) to the learned behavior or causally related to the learned behavior. I locate myself in this later camp because it appears to be a more accurate interpretation of Baldwin. I will have on another occasion to take on this tension about how to understand certain nuances with respect to the Baldwin Effect.

[20] As Dennett notes (1995: 79), “Baldwin discovered. . .that creatures capable of ‘re-inforcement learning’ not only do better individually than creatures that are entirely

selection is occurring at two levels. At one level (in the short-run), there is selection by the organism in favor of processes, which aid in molding itself (e.g., looking up while drinking water), that are part or a product of its congenital milieu; and, at a second level (in the long-run), natural selection is modifying both the existing congenital characteristics along with the newly molded modifications.[21]

Taken all together, these fortuitous improvements secured by individuals, which also serve in species preservation and phylogenetic complexity, are commonly referred to today as the "Baldwin Effect" (Richards, 1987: 451–503; Dennett, 1995: 77–80; Belew and Mitchell, 1996; Depew and Weber, 2003). [22] A formal definition is as follows:

> **Baldwin Effect**: The ability of organisms to secure advantageous behavioral variations by way of reinforced learning and mimicry to contend with problems in the short-run to better their own conditions, allowing the possibility of natural selection to retain similar advantageous variations *or* other related variations to benefit their species in the long-run.

5. THE BALDWIN EFFECT AND ALTRUISM

Having given a general account of the "Baldwin Effect," I will now explain how this theory is used to understand human altruism. According to Baldwin, human altruism is a behavior made manifest through the ontogenetic agencies of reasoning and imitation. Specifically, he argues that humans have imitative, intelligent, and social tendencies (congenital in origin; distal explanation), which can be enhanced through the imitation of behaviors of others, through explicit instruction, and through experience in general (proximal explanation). According to Baldwin (as we shall see), altruism is one such social tendency that has been modified and enhanced through various proximal mechanisms.

'hard-wired'; their species will *evolve faster* because of its greater capacity to discover design improvements in the neighborhood."

[21] Richards (1987: 482) offers the following summary of what can be called the Baldwin Effect: "Baldwin envisioned as the paradigm for his principle animals' acquiring innate behaviors similar to behaviors they originally had to learn. If a group of animals migrates into a new environment for which they lack congenital adaptations, those plastic enough to accommodate themselves through conscious learning will survive. Their ontogenetic behaviors will buffer them against the winnowing of natural selection. This safety net, according to Baldwin, will allow natural selection opportunity to accumulate chance variation that follow the path laid down by the acquired behaviors, which indeed have already been favored by selection."

[22] Emphasizing both learning and litheness with respect to niche construction and species (kin) survival, Deacon (1997: 322) defines the Baldwin Effect as follows: "Baldwin suggested that learning and behavioral flexibility can play a role in amplifying and biasing natural selection because these abilities enable individuals to modify the context of natural selection that affects their future kin." On this same point, Papineau (2004: 12) argues that the Baldwin Effect is illustrated in cases where niche construction assists an organism in bringing to fruition a set of behaviors that are only partly under genetic control. I have been, in part, defending this view in the light of Baldwin's claim that behavior can act as a supplement to one's (genetic) endowment.

Drawing on the suggestion that Sober makes in his second possibility for how human altruism came to be, the "Baldwin Effect" can be viewed as making use of the mechanism of group selection as an explanatory device. Baldwin's own words say as much with regard to the origin and maintenance of the human tendency for social behavior:

> Only so far as the individuals of a group are 'socii', members capable of cooperation and willing to cooperate with their fellows, will the group 'hold together' effectively, in competition with other groups... [T]he selective unit, considered from the external or social point of view, is a *group of individuals*, greater or smaller as the utility subserved may require (Baldwin, 1980: 43).

Baldwin explains (Baldwin, 1980: 44) that there is selection working at two levels related to the group: (1) 'intra-group' selection and (2) 'inter-group' selection. The former is social selection within the group which is geared toward socializing the individual to the group; and the latter is selection in favor of those groups that are more effectively socialized as opposed to those competing groups that are not as cohesively organized.[23] For example, human groups that are able to organize some kind of military will have a distinct advantage over those that do not. Those individuals who participate in military life quite possibly represent the altruists of the group. That is, those in the military genuinely sacrifice more for the group (greater chance of mortality) than they receive. I will turn to how Baldwin thinks such a phenomenon as altruism (e.g., voluntary military participation) is possibly sustainable.

It is from intra-group selection that altruism is maintained in the individual human organism, according to Baldwin. Through a combination of education, imitation, and voluntary cooperation from the individual, the "impulses" of sympathy and altruism (congenital tendencies) are molded as the individual develops within the group.[24] Of course, the enhancement of the "impulses" of altruism and sympathy are passed on from generation to generation (phylogenetic retention of altruistic behavior). Such transmission,

---

[23] With respect to population cohesiveness, Pagel and Mace (2004) argue that there do exist cultural adaptations, which are responses to social life and population density. Interestingly, with respect to cultural variation, they argue (2004: 277), "Humans are especially likely to copy common traits, a tendency that works to reduce variation within a culture but increase it between cultures." With respect to altruism, they go on to suggest (2004: 278) the following: "Humans have worked out a way to discard the need for [genetic] relatedness as a means of ensuring cooperation: uniquely among the animals, humans may carry a set of behavioral adaptations specific for promoting cooperation and reciprocity, even when [genetic] relatedness is low between group members. According to the doctrine known as strong reciprocity, humans are predisposed to cooperate with others, to make fair distribution of gains, and to punish those who fail to cooperate, even at a cost to themselves and with no expectation that these costs will be repaid." This seems quite in keeping with Baldwin's views about the sources of social cohesiveness. The main qualification is that migration between groups must remain low. Since this is difficult, the presence of mid to high levels of migration may very well help explain the variation of altruism both within and across the many different cultural landscapes.

[24] The neurobiology of these basic social-bond impulses are discussed by Panksepp (1998: 246–279) and Damasio (1999: 279–295), while Griffiths (1997) explores some of the philosophical issues surrounding human emotions.

in the case of soldiers in the military, could occur by the group's praise of bravery, blatant punishment, signaling/behavioral cues, individual imitation, and/or education with regard to the importance of security.[25] Baldwin refers to such a transmission as 'Social Heredity' (Baldwin, 1980: 61–67). This intra-group selection and molding of the individual is made quite clear by Baldwin:

> But we now see that the biological and individualistic sort of struggle does not represent the method of social selection. Struggle to be or to become effective and successful units or socii in an organization of self-controlled individuals-that is the proper form of statement. This involves the voluntary cooperation of individuals in the social situation... It is the person as such, the socialized self, upon whom the fittest conduct must terminate and upon whom it must originate, whether, in this case or that, it be embodied in one's own interest or in the interests of others. Sympathy and altruism are the socialized and transformed impulses of the growing individual, who is educated into a higher selfhood (Baldwin, 1980: 64-65).

Interestingly, because Social Heredity is the mechanism that sustains altruism, such an other-regarding behavior as voluntary military participation could be made manifest in a variety of ways. This is evident in Baldwin's belief that the socialized self may include elements of "self-interest" and "interests for others." Baldwin is tacitly calling for a pluralistic repertoire of motivations for human other-regarding behavior. It could be possible that an individual is taught to be other-regarding only to the extent that such other-regarding behavior enhances his reproductive success or reduces the possibility of punishment. Alternatively, an individual could simply imitate the other-regarding actions of others and be completely oblivious to any kind of cost/benefit analysis. Also possible is that a person may reflectively decide that his other-regarding actions should be either greater or lesser than the other-regarding actions of others. Lastly, a reflective individual may prefer (or be taught) to put the needs of others ahead of her own (e.g., Mother Theresa or Mahatma Gandhi). Baldwin's concept of Social Heredity allows for all of these possibilities.[26]

---

[25] Note that these various cultural ways of securing other-regarding behavior (different from or compatible with either inclusive fitness or reciprocal altruism) are examined by contemporary scholars. For example, the role of punishment as a means of securing other-regarding behavior is discussed by Yamagishi (1986); Ostrom *et al.* (1992); Fehr and Gaechter (2002), and Boyd and Richerson (1992); costly signaling is discussed by Smith and Bliege-Bird (2000) and Gintis *et al.* (2001); indirect reciprocity is explored by Sugden (1986), Nowak and Sigmund (1998), Leimar and Hammerstein (2001), Panchananthan and Boyd (2003); and cultural group selection is taken up by Gintis (2000) and Boyd *et al.* (2003).

[26] It should be clear that Baldwin is defending what many today call cultural evolution. For example, much like Boyd and Richerson (1985) and Henrich and Boyd (2001), Baldwin defends a dual- inheritance account. That is, Baldwin clearly thinks that culture is a system of inheritance much like (though not exactly like) genetic inheritance. Note, however, Boyd and Richerson are not primarily concerned with genetic fitness, but with the processes associated with how culture is transmitted within and across culture (Barkow, 1989: 253–254). Baldwin, on the other hand, takes seriously the link

The result of Social Heredity is that the group can solve short-term problems through altruism and cooperation, a strategy which promotes survival in the short-run until selection can come along and gradually fix the trait (or some other closely related trait) to ensure survival of the group. Modifications that are retained in tokens may very well ensure the survival of the type. Baldwin summarizes all of this as follows:

> [Social heredity] *keeps alive a series of functions which either are not yet, or never do become, congenital at all.* It is a means of extra-organic transmission from generation to generation. It is really a form of heredity because (1) it is a handing down of physical function; while it is not physical heredity. It is entitled to be called heredity for the further reason (2) that *it directly influences physical heredity in the way mentioned*, i.e., it keeps alive variations, thus sets the direction of ontogenetic adaptation, thereby influences the direction of the available congenital variations of the next generation, and so determines phylogenetic development. I have accordingly called it "Social Heredity" (Baldwin, 1896: 537).

In brief, Baldwin's point is that if an ontogenetic modification is of value to organisms, then we should expect that natural selection will favor genotypes that endow their possessors with the capacity to react adaptively with their surroundings in that respect or with respect to some other congenital variation. On Baldwin's account, then, human altruism, in all of its guises, reflects such adaptability.

Baldwin's account of human altruism can be summarized as follows:

(1) Altruism is a kind of social tendency made manifest through at least some individuals maintaining the integrity of the group to which they belong.
(2) As a social tendency, human altruism can be made manifest through a myriad of intentional states.
(3) A mind is required in order to ensure that a behavior such as altruism is passed on from generation to generation through teaching, punishment, imitation, and individual acceptance.
(4) A cost/benefit analysis may be under consideration, but not always.
(5) A comparison of acts may be under consideration, but not always.
(6) Genetic variation, which could be a benefit to the species as a whole, is "kept alive" as a result of (1)–(5).

In contrast to Sober, Baldwin provides an account of human altruism that does not separate so severely biological altruism and psychological altruism. Recall that Sober's first concern is that he insists that human preferences are crucial in understanding human altruism and must remain distinct from biological altruism, which is entirely understood in terms of reproductive fitness. Like Sober, Baldwin also considers the mind (conscious agency) as an important element in understanding human altruism. Unlike Sober, Baldwin considers the mind to play an active role in passing along information from generation to generation (as a possible substitute in the short-run for genetic transmission of such information). Educating a growing individual to achieve what Baldwin calls a

---

between the transmission of culture and genetic fitness along the lines of Lumsden and Wilson (1981, 1985). See Ananth (2001), Aunger (2000), Boehm (1982, 1996), Bowles (2004), Dunbar (1996), Feldman *et al.* (1996), Hammerstein (2003), Oyama (1985), and Sperber (1996) for details on some of the contemporary issues surrounding evolutionary explanations of culture.

higher socialized self-hood (Baldwin, 1980: 65) is one method (among others which were previously noted) of actualizing a latent behavior that is potentially fitness-enhancing both for the token and its type.

The second concern that Sober has with an evolutionary account of human altruism is also resolved. Baldwin's approach does not reduce the human mind to "a mere passive implementation of evolutionary imperatives," as Sober generally considers. In contrast to Sober's concern, Baldwin's theory leaves open the possibility that human altruistic behavior can be motivated by something other than purely other-regarding intentions without dismissing the possibility of purely other-regarding intentions. Points 1–5 above reflect the various psychological states that may be associated with human altruistic actions. Sober's layman or common-sense account of human psychological altruism captures but one of these many possibilities.[27]

Now, one might think that human altruism would never become genetically "hard-wired," because the survival-enhancing propensity brought about by phenotype plasticity would be lost. This concern rests on a bit of a misunderstanding of the Baldwin Effect with respect to altruism. First, it is the genetic impulse of sympathy that is hard-wired. The various behavioral and psychological expressions of this impulse are the sorts of ontogenetic modifications that can guide the production of other (or similar) variations on which selection works. It may very well turn out that a new variation (in the form of a mutation event or some other genetic chance event) could have a selective advantage—a possibility that would have remained improbable were it not for the learned altruistic behavior. Of course, it could be the case that the relevant altruistic behavior turns up as a chance genetic variation, allowing for the selection for a particular expression of the altruistic impulse. Thus, a given expression of the altruistic impulse could become hard-wired, but it does not follow that all forms of altruistic expression would become genetically assimilated. The point is that some flexibility may be lost, but this is to be expected given that adaptations are adaptations to local environments—in fact, a loss in some flexibility may still be fitness-enhancing depending on the degree of environmental perturbations.

6. CONCLUSION

The consequences of Baldwin's theory are many. First, Sober's sharp division between biological altruism and psychological altruism may not be necessary, and Sober's criterion for human psychological altruism is but one possible variation of human psychological altruism within Baldwin's framework. The various methods by which humans are able to propagate, modify, and retain behaviors in a given lifetime allow Sober's account of human psychological altruism to be subsumed under Baldwin's broader account of human biological and psychological altruistic behavior. Sober's insistence that human preferences have no bearing with regard to human fitness is simply too strong.

---

[27] To Sober's credit, he does take into account the various intentional states that accompany other-regarding and self-regarding behaviors in humans, but does not seriously address possible evolutionary implications (Sober, 1993a: 208–11). However, see Sober and Wilson (1998: 296–327) for a pluralistic defense (attacking psychological hedonism) of the evolution of psychological altruism.

Secondly, much in accordance with Sober's suggestion of how human psychological altruism could have evolved, Baldwin also thought that group-selection is a possible mode of selection for the origin and continued retention of human psychological altruism. Although Sober and Baldwin may disagree as to the details of how group-selection actually works (and it is not entirely clear how seriously Sober takes an evolutionary account of human psychological altruism), they both would consider some sense of 'group-selection' as a plausible way for explaining how such behavior and intentional states could have been sustained.

Thirdly, Baldwin correctly realized that organisms change the environment they inhabit as a result of ontogenetic modifications and in this way affect the selective pressures they (and their future generations) will countenance. That is, in order to survive in fluctuating environments an organism will have to find new ways of adapting, and these adaptive strategies in turn generate new environments which generate new changes, etc. to which the next producing offspring must confront. The result is that ecological opportunities, when taken advantage of, create new ecological opportunities for tokens and types.[28] Not only does such a feedback system create new environmental niches that can account for the tremendous variation in nature, but it also helps to explain how species are modified and possibly preserved.[29] This ability to respond to variation with respect to niches can also help to make sense of the various psychological states with respect to human psychological altruism.

Fourthly, Baldwin forces us to recognize that many organisms, especially humans, are active and interpretive creatures that are able not only to adapt to environments but, in fact, are able to construct their environments. That is, the ability to modify and improve phenotypes on the basis of learning and life experiences reflects selection in favor of adaptability. In order to adapt to a changing environment, an organism will often have to change its behavior accordingly. The ability to make such changes (within limits) requires a genetic program that allows for a particular range of behavioral adjustments. (Of course, organisms which live in extremely static environments may not have occasion, if able, to secure new ontogenetic variations.) The human organism should be viewed no differently. The human genome has produced a highly "plastic phenotype." In the human organism, the retention of such behavioral modifications, as Baldwin argued, is maintained through Social Heredity or possibly through genetic modification in the long-run (or both). Human psychological altruism, from this perspective, is a product of such environmental construction and problem-solving.

Lastly, both the Baldwin Effect and Baldwin's Social Heredity taken together narrow the gap between P1–P4 and C1 for a pluralistic understanding of human psychological altruism. *This narrowing of the gap is accomplished, if we take seriously the claim that the*

---

[28] This process of niche specialization and construction with respect to learning and evolution is discussed in Odling-Smee *et al.* (2003).

[29] No doubt, there are limitations and deleterious consequences. For example, a human trying to avoid a swarm of killer bees cannot just fly away, or human destruction of rain forests may create an environment with which future generations cannot cope. Phenotype plasticity has its limits. Moreover, many cultural influences (e.g., abuse of drugs) can have a negative effect on both individuals and the species to which they belong. On this score, blind mimicry can lead to harsh results.

capacity of an organism to respond to environmental stress during development is itself a hereditary characteristic (and such a response has an impact on the selection of other traits and affects the contouring of the environment). Human psychological altruism can be seen as such a response to particular environmental conditions. Baldwin's account, then, (1) requires a minor role to be played at the genetic level (i.e., it allows for a potential human biological altruistic behavior—the impulse of sympathy—to be made manifest), (2) explains that such a behavior was most likely retained for its usefulness, (3) upholds that transmission of human psychological altruistic behavior occurs through Social Heredity, which also allows for the possibility of other traits to be selected and retained at the genetic level, and (4) proposes that this process of ontogenetic modification and retention and phylogenetic retention of particular traits (or some other related traits) is an integral part of the natural world. The upshot here is that P3 of the argument in section 2 is true, suggesting that the gap between human/non-human biological altruism is not nearly as wide as Sober maintains.

ACKNOWLEDGMENTS

Earlier versions of this paper were presented at Oakland University and University of Wisconsin-Stevens Point. I would like to thank the participants at both universities for their critical suggestions. I would also like to pass along a special thanks to Marv Belzer, Jim Child, David Depew, Ben Dixon, Bill Harms, Carrie-Ann Kahn, Loren Lomasky, Fred D. Miller, Jr., Jaak Panksepp, Suri Ratnapala, and Elliott Sober for their willingness to discuss many of the elements in this essay. Most notably, Michael Bradie read earlier versions of this essay and offered detailed comments and much encouragement. Moreover, I very much thank the three referees of this journal for their critical feedback, which forced me to re-think portions of this paper. Finally, the majority of this essay was completed while I was a visiting scholar at the Social Philosophy & Policy Center in the Fall of 2003. I thank the Center for both its financial support and scholarly environment.

REFERENCES

Ananth, M. (2001). Critical review of Dan Sperber's explaining culture: a naturalistic approach. Philosophy of the Social Sciences 31: 563–571.

Ananthakrishnan, P.V. (1970). Consumer aspects of the ecosystems. In: K. C. Misra et al., Ecology (Study of Ecosystem), 241-88. A. H. Wheeler & Co., Allahabad.

Aunger, R. (2000). Darwinizing Culture. Oxford University Press, Oxford.

Baldwin, J.M. (1894). Imitation: a chapter in the natural history of consciousness. Mind 3: 26–55.

Baldwin, J.M. (1896). A new factor in evolution. American Naturalist 30: 441–451, 536–553.

Baldwin, J.M. (1980). Darwin and the Humanities. AMS Press, New York. (Originally published: 1909. Baltimore, Review)

Barkow, J.H. (1989). Darwin, Sex, and Status. University of Toronto Press, Toronto.

Belew, R.K. and M. Mitchell (1996). Adaptive Individuals in Evolving Populations: Models and Algorithms. Addison-Wesley, Reading.

Boehm, C. (1982). A fresh outlook on cultural selection. American Anthropologist 84: 105–124.

Boehm, C. (1996). Emergency decisions, cultural selection mechanics, and group selection. Current Anthropology 37: 763–793.

Bowles, S. (2004). Microeconomics: Behavior, Institutions, and Evolution. Princeton University Press, Princeton.

Boyd, R. and Richerson, P.J. (1985). Culture and the Evolutionary Process. University of Chicago Press, Chicago.

Boyd, R. and Richerson, P.J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethology and Sociobiology 13: 171–195.

Boyd, R., Gintis, H., Bowles, S., and Richerson, P. (2003). The evolution of altruistic punishment. Proceedings of the National Academy of Sciences of the United States of America 100: 3531–3535.

Bradie, M. (1994). The Secret Chain. State University of New York Press, Albany.

Cavalli-Sforza, L.L., Mendozzi, P., and Piazza, A. (1994). The History and Geography of Human Genes. Princeton University Press, Princeton

Colosimo, P.F. *et al.* (2004). The genetic architecture of parallel armor plate reduction in threespine sticklebacks. PLoS Biology 2: 0635–0641.

Damasio, A. (1999). The Feeling of What Happens: Body and Emotions in the Making of Consciousness. Harcourt Brace & Company, New York.

Darwin, C. (1882). The Descent of Man and Selection in Relation to Sex (2nd. ed., rev. and augm.). John Murray, London.

Darwin, C (1996). 1842 sketch and 1844 essay. In: T. F. Glick and D. Kohn (eds), On Evolution. Indianapolis, Hackett Publishing Company, Inc.

Dawkins, R. (1989). The Selfish Gene. (rev. ed.) Oxford University Press, Oxford.

Day, T., *et al.* (1994). Ecology and genetics of phenotypic plasticity: A comparison of two sticklebacks. Evolution 48: 1723–1734.

Deacon, T. (1997). The Symbolic Species. W. W. Norton, New York.

Dennett, D. (1995). Darwin's Dangerous Idea. Simon and Schuster, New York.

Dennett, D. (2003). The Baldwin effect: A crane, not a skyhook. In: D. J. Depew and B. H. Weber (eds), Evolution and Learning: The Baldwin Effect Reconsidered, 69-79. The MIT Press, Cambridge.

Depew, D.J. (2003). Baldwin and his many effects. In: D. J. Depew and B. H. Weber (eds), Evolution and Learning: The Baldwin Effect Reconsidered, 4–31. The MIT Press, Cambridge.

Depew, D.J. and Weber, B.H. (1995). Darwinism Evolving: Systems Dynamics and the Genealogy of Natural Selection. The MIT Press, Cambridge.

Depew, D.J. (2003). Evolution and Learning: The Baldwin Effect Reconsidered. The MIT Press, Cambridge.

Downes, S.M. (2003). Baldwin effects and the expansion of the explanatory repertoire in evolutionary biology. In: D. J. Depew and B. H. Weber (eds), Evolution and Learning: The Baldwin Effect Reconsidered, 33–351. The MIT Press, Cambridge.

Downing, K.L. (2004). Development and the Baldwin effect. Artificial Life 10: 39–63.

Dunbar, R. (1996). Grooming, Gossip, and the Evolution of Language. Harvard University Press, Cambridge.

Ehrman, L. and Parsons, P.A. (1976). The Genetics of Behavior. Sinauer, Sunderland.

Fehr, E. and Gaechter, S. (2002). Altruistic punishment in humans. Nature 415: 137–140.

Feldman, M.W. and Cavalli-Sforza, L.L. (1989). On the theory of evolution under genetic and cultural transmission with application to lactose absorption problem. In: M. W. Feldman (ed.), Mathematical Evolutionary Theory, 145–173. Princeton University Press, Princeton.

Feldman, M.W., *et al.* (1996). Individual versus social learning: Evolutionary analysis in a fluctuating environment. Santa Fe Institute Working Paper 96–05–031.

Galef, B.G. (1987). Evolution and learning before Thorndike: A forgotten epoch in the history of behavioral research. In: R. C. Bolles and M. D. Beecher (eds): Evolution and Learning, 39–58. Lawrence Erlbaum Associates, Hillsdale.

Gintis, H. (2000). Strong reciprocity and human sociality. Journal of Theoretical Biology 206: 169–179.

Gintis, H., Smith, E.A., and Bowles, S. (2001). Costly signaling and cooperation. Journal of Theoretical Biology 213: 103–119.

Godfrey-Smith, P. (2003). Between Baldwin skepticism and Baldwin boosterism. In: D. J. Depew and B. H. Weber (eds), Evolution and Learning: The Baldwin Effect Reconsidered, 53–67. The MIT Press, Cambridge.

Griffiths, P. (1997). What Emotions Really Are. University of Chicago Press, Chicago.

Griffiths, P. and Gray, R.D. (1994). Developmental systems and evolutionary explanation. Journal of Philosophy 91: 277–304.

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

Hamilton, W.D. (1964). The genetic evolution of social behavior. Journal of Theoretical Biology 7: 1–52.

Hamilton, W.D. (1975). Innate social aptitudes of man: An approach from evolutionary genetics. In: R. Rox, ed. Biosocial Anthropology, pp. 133–155. Malaby Pess, London.

Hammerstein, P. (2003) Genetic and Cultural Evolution of Cooperation. The MIT Press, Cambridge.

Harrison, G.A., Tanner, J.M., Pilbeam, D.R., and Baker, P.T. (1988). Human Biology: An Introduction to Human Evolution, Variation, Growth, and Adaptability. (3rd ed.). Oxford: Oxford University Press.

Henrich, J. and Boyd, R. (2001). Why people punish defectors—weak conformist transmission can stabilize costly enforcement norms in cooperative dilemmas. Journal of Theoretical Biology 208: 79–89.

Kitcher, P. (1985). Vaulting Ambition: Sociobiology and the Quest for Human Nature. The MIT Press, Cambridge.

Leimar, O. and Hammerstein, P. (2001). Evolution of cooperation through indirect reciprocity. Proceedings of the Royal Society of London. Series B, Biological Sciences 268: 2495–2501.

Lumsden, C.J. and Wilson, E.O. (1981). Genes, Mind and Culture. Harvard University Press, Cambridge.

Lumsden, C.J. and Wilson, E.O. (1985). The relation between biological and cultural evolution. Journal of Social and Biological Structures 8: 343–359.

Maynard Smith, J. (1964). Group selection and kin selection. Nature 201: 1145–1146.

Maynard Smith, J. (1982). Evolution and the Theory of Games. Cambridge University Press, New York.

Morgan, C.L. (1896). Habit and Instinct. Edward Arnold, London.

Nagel, T. (1978). Ethics as an autonomous theoretical subject. In: G. S. Stent (ed.), Morality as a Biological Phenomenon, 198–205. University of California Press, Berkeley.

Nesse, R.M. (2001). Evolution and the Capacity for Commitment. Russell Sage Foundation, New York.

Novak, M.A. and Sigmund, K. (1998). The dynamics of indirect reciprocity. Journal of Theoretical Biology 194: 561–574.

Odling-Smee, J.F., Laland, K.N., and Feldman, N.W. (2003). Niche Construction : The Neglected Process in Evolution. Princeton University Press, New Jersey.

Okasha, S. (2001). Why won't the group selection controversy go away? British Journal for the Philosophy of Science 52: 25–50.

Ostrom, E.J., Walker, J., and Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. American Political Science Review 86: 404–417.

Oyama, S. (1985). The Ontogeny of Information. Cambridge University Press, Cambridge.

Pagel, M and Mace, R. (2004). The cultural wealth of nations. Nature 428: 275–278.

Panchanathan, K. and Boyd, R. (2003). A tale of two defectors: The importance of standing for the evolution of reciprocity. Journal of Theoretical Biology 224: 115–126.

Panksepp, J. (1998). Affective Neuroscience: The Foundations of Human and Animal Emotions. Oxford University Press, New York.

Papineau, D. (2004). Social learning and the Baldwin effect. Paper read at Innateness and the Structure of the Mind, July 2003, Sheffield; available at: www.kcl.ac.uk/ip/davidpapineau/Staff/Papineau/OnlinePapers/SocLearnBald.htm

Queller, D.C. (1992). Quantitative genetics, inclusive fitness, and group selection. American Naturalist 139: 540–558.

Ramachandran, V. (2003). The Emerging Mind. Profile Books, London.

Relethford, J.H. (1997). Fundamentals of Biological Anthropology (2nd ed.). Mayfield Publishing Company, California.

Richards, R.J. (1987). Darwin and the Emergence of Evolutionary Theories of Mind and Behavior. University of Chicago Press, Chicago.

Ridley, M. (1996). Evolution (2nd ed.). Blackwell Science, Cambridge.

Ruse, M. (1998). Introduction to part VII. In: D. L. Hull and M. Ruse (eds), The Philosophy of Biology, 445–447. Oxford University Press, Oxford.

Ruse, M. and Wilson, E.O. (1986). Moral philosophy as applied science. Philosophy 61: 173–192.

Schlichting, C.D. and Pigliucci, M. (1998). Phenotypic Evolution: A Reaction Norm Perspective. Sinauer, Sunderland.

Shettleworth, S.J. (1998). Cognition, Evolution, and Behavior. Oxford University Press, New York.

Simpson, G.G. (1953). The Baldwin effect. Evolution 7: 110–117.

Skyrms, B. (1996). Evolution of the Social Contract. Cambridge University Press, New York.

Smith, E.A. and Biege-Bird, R.L. (2000). Turtle hunting and tombstone opening: Public generosity as costly signaling. Evolution and Human Behavior 21: 245–261.

Sober, E. (1988). What is evolutionary altruism? Canadian Journal of Philosophy 14: 75–99.

Sober, E. (1993a). Evolutionary altruism, psychological egoism, and morality: disentangling the phenotypes. In: M.N. Nitecki and D.V. Nitecki (eds), Evolutionary Ethics, 199–216. SUNY Press, Albany.

Sober, E. (1993b). Philosophy of Biology. Westview Press, Boulder

Sober, E. (1994). From a Biological Point of View. Cambridge University Press, Cambridge.

Sober, E. and Wilson, D.S. (1994). Reintroducing group selection to the human behavioral sciences. Behavioral and Brain Sciences 17: 585–654

Sober, E. and Wilson, D.S. (1998). Unto Others: The Evolution and Psychology of Unselfish Behavior. Harvard University Press, Cambridge.

Sperber, D. (1996). Explaining culture: A naturalistic approach. Blackwell, Cambridge.

Sugden, R. (1986). The Economics of Rights, Cooperation and Welfare. Basil Blackwell, Oxford.

Trivers, R. (1971). The evolution of reciprocal altruism. Quarterly Review of Biology 46: 35–57.

Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characteristics. Nature 150: 563–565.

Waddington, C.H. (1952). Selection of the genetic basis for an acquired character. Nature 169: 278.

Waddington, C.H. (1953a). Genetic assimilation of an acquired character. Evolution 7: 118–126.

Waddington, C.H. (1953b). The Baldwin effect, genetic assimilation, and homeostasis. Evolution 7: 386–387.

Waddington, C.H. (1961) Genetic assimilation. Advanced Genetics 10: 257–290.

Wade, M.J. (1980). Kin selection: Its components. Science 210: 665–667.

West-Eberhard, M.J. (2003). Developmental Plasticity and Evolution. Oxford University Press, Oxford.

Williams, G.C. (1966). Adaptation and Natural Selection. Princeton University Press, Princeton.

Wilson, D.S. and Dugatkin, L. (1992). Altruism: contemporary debates. In E. F. Keller and E. A. Lloyd (eds), Key Words in Evolutionary Biology, 29–33. Harvard University Press, Cambridge.

Wilson, E.O. (1975). Sociobiology: The New Synthesis. Harvard University Press, Cambridge.

Wynne-Edwards, V.C. (1962). Animal Dispersion in Relation to Social Behavior. Oliver and Boyd, Edinburgh.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. Journal of Personality and Social Psychology 51: 110–116.