

**Altruism, Egoism, or Neither: A Cognitive-Efficiency-Based Evolutionary Biological  
Perspective on Helping Behavior**

(Forthcoming in *Studies in History and Philosophy of Biological and Biomedical Sciences*)

Armin Schulz

Department of Philosophy

University of Kansas

3083 Wescoe Hall

Lawrence, KS 66045

[awschulz@ku.edu](mailto:awschulz@ku.edu)

<http://people.ku.edu/~a382s825/>

## **Abstract**

I argue for differences in the cognitive efficiency of different psychologies underlying helping behavior, and present an account of the adaptive pressures that result from these differences. Specifically, I argue that organisms often face pressure to move away from only being egoistically motivated to help: non-egoistic organisms are often able to determine how to help other organisms more quickly and with less recourse to costly cognitive resources like concentration and attention. Furthermore, I also argue that, while these pressures away from pure egoism can lead to the evolution of altruists, they can also lead to the evolution of reciprocation-focused behaviorist helpers or even of reflex-driven helpers (who are neither altruists nor egoists). In this way, I seek to broaden the set of considerations typically taken into account when assessing the evolution of the psychology of helping behavior—which tend to be restricted to matters of reliability—and also try to make clearer the role of evolutionary biological considerations in the discussion of this apparently straightforwardly psychological phenomenon.

## **Altruism, Egoism, or Neither: A Cognitive-Efficiency-Based Evolutionary Biological Perspective on Helping Behavior**

### **I. Introduction**

What motivates an organism to help another is still an open question, despite being quite widely discussed (see e.g. Stich et al., 2010; Davidic et al., 2006; Fehr & Gaechter, 2000; Batson, 1991; Nagel, 1970). Given this lack of a settled account of the psychological structures underwriting helping behavior, it is perhaps unsurprising that researchers have looked for new ways to investigate this issue. Among these new approaches is an evolutionary biological one: specifically, a number of authors have tried to assess the evolutionary pressures on different cognitive architectures with a view to their ability to lead to helping behavior (Sober & Wilson, 1998; see also Stich, 2007; Schulz, 2011a; Kitcher, 2011; Clavier & Chapuisat, 2013). It is this evolutionary biological take on the psychology of helping behavior that is the focus of this essay.

Specifically, I here argue that, in evolutionary biological investigations of the psychology of helping behavior, we ought to move away from just considering the *reliability* of different mind designs to lead organisms to help others—which is what the existing analyses have tended to concentrate on—and instead consider how *cognitively efficient* different mind designs are at helping others. In particular, I show that there is a certain kind of cognitive inefficiency that can characterize purely egoistically motivated organisms, and which can push populations of organisms away from featuring mostly purely egoistic helpers. This cognitive inefficiency stems from the fact that, as compared to non-egoistic helpers, egoistic helpers sometimes are bound to make helping decisions more slowly, and with more recourse to costly cognitive resources like concentration and attention.

The paper is structured as follows. In section II, I make clear how I propose to understand some of the key terms of the debate here, and set out the (somewhat modest) role I see evolutionary biology as playing in it. In section III, I discuss the existing, reliability-focused evolutionary biological analyses of the psychology of helping behavior. In section IV, I present the core, cognitive-efficiency-based evolutionary argument against psychological egoism. An extreme case of this argument leading to completely non-representational helping decisions is presented in section V. I conclude in section VI.

## **II. The Psychology of Helping Behavior and the Role of Evolutionary Biology in its Investigation**

In what follows, I understand an organism to be a *psychological altruist* if and only if it holds ultimate desires for the well being of other organisms, and a (pure) *psychological egoist* if and only if it holds ultimate desires for its own well being only (see also Sober & Wilson, 1998; Stich et al., 2010). A few points are important to note about this way of understanding altruism and egoism.<sup>1</sup>

Firstly, I here leave it open exactly what ‘well being’ consists in (Stich, 2007). The only assumption that I do make concerning this is that well being is at least correlated with fitness: increasing an organism’s well being will tend to increase that organism’s fitness (Brown et al.,

---

<sup>1</sup> Note that other understandings of these notions are possible (Garson, forthcoming). However, the ones in the text are the most widely accepted ones, and at any rate, I do not think that much of substance in this paper hangs on this particular characterization of the terms—with different characterizations, the arguments of this paper might have to be reformulated, but would remain substantively the same. Note also that this definition of egoism fits quite naturally to some recent work in evolutionary theory: for example, Grafen (1999) has argued that all organismic behaviors can be seen as the result of an organism maximizing its inclusive fitness. Finally, note that psychological altruism as defined here must not be conflated with *evolutionary altruism*: organismic traits that provide (relative or absolute) fitness benefits to other organisms (Sober & Wilson, 1998; Okasha, 2006). The latter raises different issues from the ones at stake here, and will not be discussed further in what follows. In line with this, all unqualified references to ‘altruism’ or ‘egoism’ in this paper should be taken to refer to the psychological varieties of these two theses only.

2012, 235-236). This assumption is quite uncontroversial, though, and is shared by most of the rest of the literature on this topic (see e.g. Sober & Wilson, 1998; Buller, 2005).

Secondly, psychological altruism and egoism as understood here can be easily situated within the currently popular two-systems models of the mind (Epstein et al., 1992; Sloman, 1996; Chaiken & Trope, 1999, Stanovich & West, 2000).<sup>2</sup> Two-systems models of the mind comprise two parts: an associative, reflex-based system and a representationalist, deliberative system. While there are still many open questions concerning how to best characterize the two systems and their interrelations (Chaiken & Trope, 1999; Kahneman, 2003; Campbell & Kumar, 2012), for present purposes, it is enough to make the uncontroversial point that the two systems differ in how they generate the organism's behaviors.<sup>3</sup>

In the case of the reflex-based system ('system 1' in what follows), the action is the result of a mapping—i.e. a look-up table—between states of the world (as they are perceived by the organism) and actions. In other words, system 1-behaviors are to be seen as *reactions* to the perception of a state of the world, rather than as the result of a genuine practical inference (see also Schulz, 2013).

By contrast, in the case of the deliberative system ('system 2' in what follows), action generation is mediated by content-bearing—i.e. representational—mental states.<sup>4</sup> In particular, the organism is taken to form or have, on the one hand, representations about what the world is

---

<sup>2</sup> Two-systems models of the mind—while not without their detractors (Glimcher et al., 2005; Kruglanski & Gigerenzer, 2011)—are among the most widely accepted models of cognitive architecture currently in the literature, and have significant empirical and theoretical support (Epstein et al., 1992; Stanovich & West, 2000; Haidt, 2001; Kahneman, 2003). For this reason, the focus on them should not be seen as a greatly restrictive assumption. At any rate, none of the conclusions of this paper actually hang on the details of these models, and could also be formulated with reference to alternative models (such as the heuristics-focused view of Kruglanski & Gigerenzer, 2011).

<sup>3</sup> Note that it is also possible that both systems are active at the same time, and attempt to influence the organism's behavior (Greene, 2008; Kahneman, 2003; Haidt, 2001); fortunately, nothing in what follows concerns or depends on how these conflicts are resolved.

<sup>4</sup> While there is a lot of controversy over what it takes for a mental state to be content-bearing (Millikan, 1984, 2002; Papineau, 1987; Dretske, 1988; Fodor, 1990; Prinz, 2002), for present purposes, this can be left open.

like ('beliefs') and, on the other, representations about what the world ought to be like ('desires'); it is then further taken to combine these to make a decision about what to do. Importantly, within this representational practical inference, a distinction between two kinds of desires can be made: *ultimate* and *instrumental* ones. While the exact details of this distinction are controversial (see e.g. Goldman, 1970), for present purposes, it is enough to see instrumental desires as resulting from a piece of deliberation based on other desires and beliefs, and ultimate desires as desires that an organism has that are not the result of this kind of deliberation (see also Stich, 2007). It is in this sense that the term 'ultimate desire' in the definitions of 'psychological altruism' and 'psychological egoism' is to be understood here.

Thirdly, as understood here, altruism is a pluralist mind design, while (pure) egoism is a monist one (Sober & Wilson, 1998). Altruism does not require that all of an organism's ultimate desires are for the well being of others; only that some of them are. Indeed, to the extent that an organism has desires at all, it is reasonable to assume that it has ultimate desires for increasing its own well being. What this further implies is that there will (sometimes) be a major difference between the altruist and the egoist in the way they make decisions about whether to help someone else (this will only sometimes be the case, as an organism need not be an altruist about helping everyone else). The egoist will always *reason* about this—she will decide whether to help by determining whether helping will contribute to her own well being. By contrast, the altruist will use a non-reasoning-based process to determine which of its ultimate goals to pursue: different situations will 'trigger' different ones of her ultimate desires to be the determinant of her actions. Since the altruist has more than one ultimate desire, the decision between the latter has to be based on a non-reasoning-based process. Figures 1 and 2 illustrate this difference.

[Figure 1]

[Figure 2]

The fourth and final point to note concerning psychological altruism and egoism as understood here is that they are not exhaustive of the space of possibilities. In particular, there are two main ways in which an organism might be neither an altruist nor an egoist. On the one hand, an organism might be (partly) driven by ultimate desires that are neither for their own well being nor for that of some other organism. For example, an organism might have an ultimate desire to make works of art—which concerns neither the organism’s own well being, nor that of other organisms’ (though it might impact one or both of these), and so is neither altruistic nor egoistic. In what follows, I will call organisms of this type “behaviorist helpers”. On the other hand, an organism might, in the relevant circumstances, be driven by reflexes only (Mackintosh, 1994; Dickinson & Dyer, 1996; Grau, 2002; Kacelnik, 2012). Since their behavior is not then based on ultimate desires at all, they are acting neither altruistically nor egoistically (even though the outcome of their behavior might well be an increase in their or some other organism’s well being). In what follows, I will call organisms of this type “reflexive helpers”.

With the content of the theses of psychological altruism and egoism thus clarified, it next needs to be noted that it is still controversial which organisms should be seen to be (pure) egoists, which altruists, and which behaviorist or reflexive helpers (Stich et al. 2010). It is for this reason that considering an evolutionary biological perspective might seem tempting here: it may offer a way to add considerations to the ones typically taken into account, and thus to make some progress in settling this issue.

However, this then raises the question of what role such an evolutionary biological perspective can play in the dispute surrounding psychological altruism and egoism. The main worry in this regard is that determining the evolutionary pressures impinging on a given trait is hard enough when the trait is known to exist and relatively easy to investigate empirically; it is likely to be much harder for traits—like psychological altruism or egoism—whose existence is still controversial and the investigation of which is empirically difficult (Brandon, 1990; Buller, 2005; Richardson, 2007; Sober, 2008). Given this, one may be inclined to dismiss evolutionary biological analyses of the psychology of helping behavior as being unable to add much to the debate surrounding the latter (see also Stich, forthcoming). However, there is some reason to think that evolutionary biological considerations can be quite illuminating here after all.

To see this, begin by considering some of the main challenges faced by evolutionary biological analyses in general (Brandon, 1990; Sober, 2008). As part of this kind of analysis, it is typically necessary to estimate: (1) the relative importance of different evolutionary determinants (e.g. natural selection and drift), (2) the details of the way in which the relevant evolutionary determinants work (e.g. what natural selection selects for and why), (3) the developmental and epigenetic factors constraining the operation of the relevant evolutionary determinants (e.g. the existence of pleiotropic effects or developmental bottlenecks), and (4) the ancestral conditions of and the available variation in the traits in question.

Now, it is indeed plausible that addressing these points in detail and with significant confidence is hard for traits like psychological altruism: establishing the relative importance and detailed workings of different evolutionary determinants is difficult for traits that do not fossilize and which are only tenuously connected to easily measureable traits (like morphological structures or artifacts) (Richardson, 2007; though see also Mithen, 1990; Sterelny 2003, 2012).



However, this does not mean that it must always remain impossible to say anything at all concerning (1)-(4) when it comes to the psychology of helping behavior.

In particular, it can be noted concerning point (1) that, since helping behavior is likely to be adaptively very important, it is plausible that natural selection played a relatively large role in the evolution of the ways in which an organism decides to engage (or not) in this kind of behavior (Sterelny, 2003, 2012; Sober & Wilson, 1998). In relation to point (2), it can be noted that much can be said about the details of the way in which natural selection operated in the context of the psychology of helping behavior—indeed, pointing this out is the aim of the rest of this essay. Finally, in relation to points (3) and (4), it can be noted that, due to the fact that the different ways of deciding to help others can be seen as instantiations of the same two-systems model of the mind, it is at least not a completely implausible starting assumption that: (i) there are few constraints that do not fall equally on all of the different ways of deciding to help others; (ii) the ancestral condition is unlikely to have significantly biased the path of the subsequent evolution of the motivation to help others; and (iii) the relevant variability in altruistic, egoistic, and other kinds of helping behaviors was present (see also Sober & Wilson, 1998).

Now, it is true that these answers to questions (1)-(4) are far from established with a great degree of certainty. Therefore, it is also true that any resulting hypotheses concerning the evolution of the psychology of helping behavior cannot count as having been highly confirmed (Richardson, 2007; Buller, 2005). However, for two reasons, this need not be completely damning to the evolutionary biological analysis of psychological altruism and egoism.

Firstly, there is little a priori reason to think that, over time, we cannot get more detailed or better supported answers to questions (1)-(4) (despite what is claimed in Lewontin, 1998). For example, advances in cognitive and neural ethology might enable us to determine with more

certainty which organisms (if any) are currently altruistically motivated, which are egoistically motivated, and which are neither. In turn, these kinds of insights can then be used as the basis for a comparative study of the evolution of the psychology of helping behavior that can further confirm or disconfirm the present analysis.

Secondly, if the evolutionary biological perspective is only used to provide *evidence* for—and not considerations that are meant to lead us to *accept*—hypotheses about the frequency of altruists, egoists, and other kinds of helpers in a population of organisms, the above answers to (1)-(4) can be sufficient (Sober & Wilson, 1998; Sober, 2008; Schulz, 2011a, 2013).<sup>5</sup> Put differently: if the goal is merely to add some further considerations to the ones already available concerning the existence of psychological altruism, egoism, and other forms of helping dispositions (from psychology, neuroscience, economics, philosophy, etc.), then the fact that we are not entirely sure about all aspects of the evolutionary conditions driving these traits is not so problematic. Of course, this then implies that the evolutionary biological perspective does not hold a particularly central position in the debate surrounding the psychology of helping behavior—but this can be easily accepted, and does not mean that it has no role to play in this debate at all.

### **III. Altruism, Egoism, and Reliability**

To date, the standard account of the evolution of psychological altruism is based on the reliability with which the latter leads to adaptive behavior. In relatively rough form, the idea here is as follows (for more detail, see Sober & Wilson, 1998; Stich, 2007; Schulz, 2011a).

---

<sup>5</sup> Another use of evolutionary biological considerations in psychological disputes is purely heuristic (Machery, forthcoming; Schulz, 2011a, 2012). However, since this is not central here, I will not discuss this further.

Consider a case where it is *adaptive* for an organism O to help another organism S—e.g. because S is O's offspring, or because doing so would induce S to reciprocate in the future, thus reducing O's risk of being left helpless. (Here and in what follows, I use 'O helps S' as short for 'O increases S's well being'.) Given this, for O to indeed decide to help S, O needs to (a) recognize that S is in need of help, (b) know what kind of help S needs, (c) be able to provide this kind of help, and (d) be motivated to provide this kind of help. So, how can (a)-(d) be accomplished? More generally, how can it be ensured that (a)-(d) are *reliably* the case?

Now, when it comes to (a)-(c), it is necessary to harmonize the needs of S and the perceptions, beliefs, and abilities of O. In general, this should not be taken for granted or assumed to be easily accomplished, but in at least some cases this kind of harmonization does seem possible to build up and maintain. Because of this, it is sufficient for present purposes to simply assume that (a)-(c) are satisfied, and focus on point (d): the question at stake here is just the one of how O can be reliably *motivated* to help S.

When it comes to (d), though, things are complex. In particular, assuming that O decides if and how to help S using representational mental states (an assumption to which I return below), two options are typically considered for how O might decide to help S: O could be an egoist and believe that increasing S's well being is most conducive to increasing its own well being (as in figure 1), or O could be an altruist and have an ultimate desire to increase the well being of S (as in figure 2).

It further has been argued that the altruistic solution is likely to be more reliable in leading O to help S—and thus (by assumption) more reliable in getting O to act adaptively (Sober & Wilson, 1998). Primarily, this is because there is less that can happen that would lead to O not being motivated to help S. In the egoistic case, there is a chance that O starts to believe that

doing something other than increasing S's well being is most conducive to its (i.e. O's) well being. However, if O does change its beliefs in this way, then it is bound to start acting maladaptively, since in this case, increasing S's well being is (by assumption) adaptive. None of this, though, can happen if O is an altruist: for then, it will always be motivated to help S.<sup>6</sup>

However, while cogent as such, it has recently been suggested that this argument might ultimately underestimate the potential reliability of purely egoistic architectures (Stich, 2007). This is due to the fact that, in order for a purely egoistic organism O to be a reliable helper of S, it is sufficient for O's help-inducing cognitive state (viz., that the best way for it—i.e. O—to increase its well being is to increase S's well being) to be 'sub-doxastic' (Stich, 2007). Sub-doxastic states are mental states that function much like beliefs, but which are more rigid than beliefs, in that they are not amenable to updating and other forms of change—i.e. they are fairly fixed components of an organism's mental life. Importantly, these states have also been appealed to in many other contexts, and their existence is quite widely accepted by now (Fodor, 1983; Carey & Spelke, 1996; Stich, 2007). Now, while there is more that can be said about these states, what matters most for present purposes is that they are immune to the worries about maladaptive belief changes mentioned above: after all, what makes these states unique is precisely the fact that they *cannot* be changed easily. Given this, it seems that there is no reason to think that a sub-doxastically motivated egoist either could not exist or that it would not be just as reliable to help others as an altruist would be.

Now, it may be possible to respond to Stich's (2007) argument here (Schulz, 2011a).

Fortunately, for present purposes, it is not necessary to assess the exact extent (if any) to which

---

<sup>6</sup> Since altruists are pluralistically motivated, it is also the case that they can still be motivated to help S even if one of their motivational systems gets damaged (Sober & Wilson, 1998). I shall not discuss this further here, though, as doing so would not affect the overall conclusion reached—namely, that there are also cognitive-efficiency-based considerations that impact the evolution of altruism or egoism.

an altruist is more reliable than an egoist in helping another organism. The reason for this is that, as I try to make clearer in what follows, there is also another evolutionary pressure—based on considerations of cognitive efficiency—that ought to be taken into account here. Whether this other evolutionary pressure is the only driver of the evolution of the psychology of helping behavior or whether it acted in concert with the reliability-based adaptive pressures discussed above can be left for another occasion (especially when keeping in mind the modest, evidential goals of the present inquiry).

#### **IV. Altruism, Egoism, and Reciprocation-Focused Behaviorist Helping**

The key idea to be defended in what follows is that considerations of cognitive efficiency—i.e. assessments based on which cognitive systems are faster and require fewer cognitive resources such as concentration and attention for their functioning—suggest that, when it comes to the generation of helping behavior, there can be something evolutionary unstable about purely egoistic motivational architectures. To understand this argument, it is necessary to start by noting the evolutionary pressures on representational, system 2-based decision making in general (Godfrey-Smith, 1996; McFarland, 1996; Dickinson & Balleine, 2000; Millikan, 2002; Sterelny, 2003; Schulz, 2011b, 2013). Here, the following two broad points can be noted.

Firstly, it is now widely accepted that representational decision making comes with costs relative to reflex-based and other non-representational forms of decision making (Clark, 1997; Millikan, 2002; Sterelny, 2001; 2003; Schulz, 2011b, 2013). These costs can take a number of different forms, but are widely thought to include (i) increased demands for cognitive resources like concentration and attention and (ii) decreased decision making speed (Clark, 1997; Haidt, 2001; Stanovich & West, 2000; Schulz, 2013). Note that these costs are cognitive, and not

energetic: the point is not that relying on more rather than fewer representational inferences (by itself) increases the organism's metabolic demands, but rather that it increases its demands for cognitive resources like concentration, attention, and decision making time.<sup>7</sup> Note also that it is *not* generally accepted that representational decision making is always more error prone—and thus less reliable—than reflex-based decision making. In fact, the reliability of the two systems depends on the details of the case (see also Schulz, 2013).

However, secondly, representational decision making also comes with a distinctive set of benefits. Now, it is controversial exactly what these benefits are—however, it is generally agreed that they turn on the fact that this kind of decision making relies on (practical and theoretical) inference making. Furthermore, three main accounts have been proposed of why the latter can be adaptive. Firstly, some have argued that the ability to make inferences can be beneficial, as it allows an organism to make decisions that it could not make otherwise—e.g. because these decisions require an understanding of (parts of) the (causal) structure of the world (Dickinson & Balleine, 2000; Millikan, 2002; Papineau, 2003; Dayan, 2012). Secondly, some have argued that the ability to make inferences can be beneficial, as it allows an organism to fine-tune its behavioral responses to the environment—e.g. because it enables an organism to react differently even to (close to) the same situation (Sterelny, 2003). Finally, some have argued that making inferences can be beneficial, as it can simplify the decision situation of the organism in specific ways—in particular, it can enable the organism to save memory by placing a given decision into a pattern of similar other decisions (Schulz, 2011b; 2013).

Fortunately, for present purposes, it is not necessary to commit to one of these particular accounts of the benefits of representational decision making. The reason for this is that there is a

---

<sup>7</sup> Of course, this is not to say that representational decision making *could not* be more energetically costly than other forms of decision making. The point is just that it *need not* be.

certain kind of case in which, according to all of the above accounts, representational decision making is *not* likely to be beneficial. This kind of case concerns “separable” decision situations—i.e. decision situations that cannot be handled more easily by placing them in the context of other decision situations—in which, by and large, the same organismal response is adaptive. This can be seen easily from the fact that, in this kind of case, the organism is unlikely to gain much by inferring its response to the situation, but would still have to pay the cognitive costs of doing so (i.e. loss of decision making speed and increased demand for cognitive resources). In a bit more detail, this point can be spelled out as follows.

Assume an organism O faces many instances of a type of decision situation E. Assume also that the set of factors that determines the most adaptive response to any particular instance of E is unique to E, so that there is little to be gained by placing this decision situation in the context of other decision situations (i.e. assume that decision making in E is “separable” from decision making in other situations). Assume next that most of the instances of E feature the same causal setup and do not reward novel or flexible responding; instead, the same organismal response R is adaptive in all of them. Finally, assume that in those instances of E where the causal setup is changed or where novel or flexible responding would be adaptive, doing R does not lead to major losses in fitness (relative to the most adaptive response). Now, in a case like that, there is little that an organism can gain by inferring the best response to the instance of the situation it is facing, instead of simply always doing R. In particular, it would lose time and cognitive resources from making this inference, but only gain the avoidance of the mistake of producing R when it would not be adaptive to do so. Since that mistake was assumed to be rare and not very costly, though, the overall fitness balance is likely to be negative here. In short: if an organism faces many instances of a “separable” decision situation that, by and large, do not causally differ

from each other and which do not require novel or flexible responding, representational reasoning is unlikely to be adaptive.

This last point is crucial for the discussion surrounding the psychology of helping behavior, for it turns out that many helping situations are exactly of this kind—i.e. many helping decisions are best approached by avoiding any kind of representational inference towards the desire to help. This matters, as the adaptiveness of avoiding this representational inference speaks in favor of the evolution of non-egoistic ways of making some helping decisions: after all, what defines altruists and “behaviorist helpers” is precisely the fact that their desire to help is *triggered* by the perception of another organism in need, rather than *inferred* from a more basic other ultimate desire. In more detail, this idea can be spelled out as follows.

There are two important kinds helping situations that, for many organisms, reward the—at least partial—avoidance of representational decision making. These situations are (i) cases where an organism needs to decide whether to provide help to its offspring, and (ii) cases where an organism needs to decide whether to help those who have helped it in the past. (There may be other helping situations that could be cited here, too, but these two are particularly clear and important. At any rate, the existence of these two cases is enough for the rest of the argument to go through.) Consider these two cases in turn.

When it comes to decisions about whether to help one’s offspring, note firstly that the well being—and thus, by assumption, the fitness—of an organism’s dependent children is closely related to an organism’s own fitness. As commonly defined, fitness is an increasing (probabilistic) function of the (expected) number of offspring—including grand-offspring—an organism has (Sober, 2001). Hence, increasing the well being (fitness) of an organism’s offspring generally increases the organism’s own well being (fitness) as well: the higher the well



being (fitness) of an organism O's offspring, the more offspring that offspring is likely to have—and thus, the more grand-offspring O is likely to have.

Secondly, note that it is often adaptive for an organism to *itself* take steps to ensure that the well being (fitness) of its offspring is high—i.e. to consistently be disposed to help that offspring. A parent is often best placed to increase the well being of its offspring—it knows best how to do so, and often has the best means at its disposal to achieve this. Moreover, for many organisms, there is little else that competes with the importance of caring for their offspring (at least when that offspring is young): specifically, it is a widely known fact about many mammals that the parent-offspring—and especially the mother-offspring—bond is exceptionally deep and of great importance for both the parent and the offspring (Chevrund & Wolf, 2009; Thometz et al., 2014; Curley & Keverne, 2005; Churchland, 2011).

In turn, this suggests that, for many organisms—including humans—deciding to help their offspring is precisely one of the cases where representational decision making is unlikely to be adaptive. Specifically, given the above, it is the case that, for many organisms, (a) the question of whether to help their offspring is “separable” from other helping-related questions—i.e. it is a decision that can respond to a unique set of factors: namely, whether the organism in need of help is an offspring (Curley & Kaverne, 2005; Churchland, 2011). Furthermore, it is the case that (b) it is very often adaptive for them to help their offspring (Thometz et al., 2014), and (c) not helping their offspring when it would be adaptive to do so is more adaptively costly than helping them when it would not be adaptive to do so (Curley & Kaverne, 2005). Given the fact that inferring whether to help their offspring comes with costs, it is therefore the case that many organisms are well placed to decide to help their offspring by relying on non-representational decision making: for figuring out whether it is adaptive to help their offspring in every particular

circumstance is generally not worth the effort as compared to simply presupposing it is (Brown et al., 2012).

Note that the issue here just concerns the question of *whether* the organism is to help its offspring, and does not concern the question of *how* the organism is to (best) help its offspring. The latter question might reward representational decision making—i.e. it might not fall into the set of “separable” decision situations that do not reward causal reasoning or flexible responding. The point made here, though, is that the initial decision concerning whether to help the offspring at all is best taken non-representationally. (Alternatively, if one prefers to see the decision about whether to help offspring and the decision about how to help the offspring as parts of the same overarching decision, the point made here is that it is adaptive for the *first step* of that decision to be non-representationally triggered, rather than representationally inferred.)

This matters, for precisely this is what altruists do: instead of *reasoning* about whether to help their offspring, they let their perception or belief that their offspring is in need of help simply trigger their ultimate desire to help that offspring. In this way, it becomes clear that many organisms—i.e. those that frequently enough face strong adaptive pressures to help their offspring—are most adaptively structured if they are altruists: for they then avoid paying the costs of representationally deciding whether to help their offspring, and forgo few benefits in the process. Three points are important to note about this argument.

Firstly, the above argument is not premised on the claim that it is always adaptive for all organisms to help their offspring; indeed, the latter is known to be false (Trivers, 1974). The point is just that, for a significant number of organisms, helping their offspring is sufficiently often the adaptive response to have the disposition to help their offspring stem from an altruistic

motivational structure be adaptive. Importantly, moreover, it is plausible that humans are among these organisms: humans generally fit conditions (a)-(c) above (Churchland, 2011).<sup>8</sup>

The second point to keep in mind concerning the above argument is that it turns on the *cognitive* costs that come from representational decision making, and not on *energetic* ones. This is important to note, as it shows why the perspective developed here is in fact consistent with that in Sober & Wilson (1998). While the latter argue that there is no reason to think that psychological altruism is any more or less energetically costly than psychological egoism (Sober & Wilson, 1998, p. 322), this does not speak to what is true when it comes to the cognitive efficiency of the two cognitive architectures—and it is precisely the latter that is relevant here.<sup>9</sup> Relatedly, recall also that I am here not presuming that either altruism or egoism is more *reliable* in leading the organism to help. My point is that the mere fact that the egoist relies more on representational decision making sometimes speaks *against* its adaptiveness relative to that of the altruist: for this increased reliance on representational decision making comes with increased needs for cognitive resources such as time, concentration, and attention that (sometimes) are not compensated by major benefits.

---

<sup>8</sup> It is known that infanticide is not uncommon among humans (Hausfater & Hrdy, 1984; Hrdy, 1999). However, this is consistent with the account defended in the text. On the one hand, even if humans have an ultimate desire to help their offspring, this does not mean that it is always possible or straightforward to act on this desire: e.g. when several offspring are in need and resources are scarce, complex decisions about how to best allocate the available resources among the different offspring need to be made (this may include deciding to let one of the offspring starve). On the other hand, it is also possible that different human populations faced sufficiently different environmental conditions, and only some found altruism about their offspring to be adaptive (see also Penke, 2010). For present purposes, it is enough to note that it is plausible that at least some human populations are among those organisms that have found altruism to be adaptive.

<sup>9</sup> Of course, it is possible that, at least sometimes, the cognitive efficiency of an organism can be indirectly measured by its caloric intake—a more cognitively efficient organism might have more time foraging and consuming food, for example. However, it would then remain the case that the more cognitively efficient mind design might itself be just as energy hungry as the less cognitively efficient one—it is just that the former can cause the organism to lower its energy consumption overall. In other words: it may be that developing and maintaining traits T and T' is energetically equally costly, but that, once either T or T' is in place, T leads the organism to lower its overall energy consumption relative to T'. I thank Justin Garson for useful discussion of this point.

The third and last point to note concerning the above argument concerns the question of whether the costs that come from making an inference towards the desire to increase an offspring's well being are large enough to put significant amounts of adaptive pressure on egoists to become altruists. In particular, given that much inference making might remain even after the shift to altruism (namely from the ultimate desire to increase the offspring's well being to the exact form of helping behavior that O ought to engage in), one may wonder if avoiding the inference to the desire to increase the offspring's well being indeed yields major adaptive benefits for O.

In response, two remarks can be made. On the one hand, even if the costs that come from making the inference to the desire to increase an offspring's well being are not major, this does not mean that these are not costs that the organism would be better off avoiding. In particular, as is well known, evolution by natural selection does not need major fitness benefits to be highly effective (Haldane, 1932; Dawkins, 1986). On the other hand, it is important to keep in mind the goals of the present evolutionary biological analysis. The argument here is meant to provide *evidence*—a reason—for the evolution away from (pure) psychological egoism; the goal was not to show that altruistic cognitive architectures *definitely evolved*. While the latter goal is indeed not attainable by the above argument, the former is—even with relatively minor adaptive pressures on egoists (Sober & Wilson, 1998).

Importantly, though, this is not where the discussion should be seen to end—for there is a very similar argument that can be given concerning a second type of helping situation: namely, cases where an organism needs to make decisions about whether to help an organism that has helped it in the past (i.e. decisions concerning reciprocation—Trivers, 1971; Sachs et al., 2004). In particular, for a number of organisms, it is plausible that (a) the question of whether to help

other organisms that have provided aid in the past is a “separable” problem (i.e. it is one that can be answered by just considering whether the organism in need of help has indeed provided help in the past), and (b) a positive answer to this question is sufficiently often adaptive, to make it that case that (c) it is adaptive to let the desire to reciprocate past help simply be *triggered*—instead of *inferred*—from a perception or belief that a reciprocator is in need. In more detail, the idea here is as follows.

Firstly, the question of whether to help an organism that has helped one in the past (and / or which is in a position to reciprocate that help in the future) is often a problem that is best approached within the confines of the concept of reciprocation only (Sachs et al., 2004; Sterelny, 2003, 53-54). Further, there is no doubt that the gains of taking part in reciprocal interactions are often large—e.g. they can make the difference between life and death (Carter & Wilkinson, 2013). Finally, the costs from unnecessarily reciprocating past help are often small—they are restricted in time and space, and are often of minor importance to the helper (Sachs et al., 2004). In short: for many organisms—including humans (Fehr & Gaechter, 2000)—reciprocation is “separable” and generally adaptive.<sup>10</sup>

In this way, by exactly the same reasoning as above, it becomes clear that an organism for which the desire to reciprocate past help is an ultimate one would seem to be more cognitively efficient than an egoist that always has to make an inference to that desire (assuming that, for this organism, reciprocation is indeed “separable” and sufficiently often adaptive). Note again that, for this argument to go through, it need not be the case that reciprocation is literally always

---

<sup>10</sup> Note that being a reciprocal responder is not straightforward—which may explain why there are very few confirmed cases of reciprocation in nature, despite the fact that the latter would often seem to be adaptive (Hammerstein, 2003; Clutton-Brock, 2009; though see also Carter & Wilkinson, 2013). In particular, reciprocation requires keeping track of who helped whom in the past, and then relating that knowledge to current interactions (not to mention the fact that the partners in the interaction have to interact multiple times). This is cognitively complex. Fortunately, there is no doubt that at least some organisms—including, especially, humans—are able to do this (see e.g. Fehr & Gaechter, 2000; Carter & Wilkinson, 2013). Thus, while the upshot of the argument in the text is once more restricted in its reach—it only applies to some organisms—this does not mean it has no value whatsoever.

adaptive. The point is just that it is plausible that there are cases in which the benefits from treating reciprocation as always worth doing—namely, saving the computational costs of determining whether it is in fact adaptive on this particular occasion—outweigh the costs from doing so—namely, at times acting maladaptively. In this way, it becomes clear that some egoists (i.e. those that face strong adaptive pressures to be reciprocators) will face adaptive pressures to become organisms driven by an ultimate desire to reciprocate past help.

At this point, it is important to note that this ultimate desire to reciprocate past help is, in one important way, different from the altruist's desire to help their offspring: it is not targeted at any particular organism (such as offspring), but it is targeted at a specific kind of behavior (reciprocation). This matters, since organisms (partly) driven by ultimate desires for particular forms of behavior are, as made clearer above, neither altruistic nor egoistic. Instead, they are “behaviorist helpers”: organisms with a strong commitment to a certain kind of behavior. While this behavior may be directed at others, and while it might result in increasing their well being, the latter is not what the desire is about. So, while reciprocation-focused “behaviorist helpers” are not egoists, they are not altruists either. For this reason, the reciprocation-focused case leads to the evolution away from egoism, but not towards altruism.<sup>11</sup>

Summing up the argument of this section, this all yields the following. I have tried to show that there are evolutionary pressures for many organisms to move away from being purely egoistically motivated—either towards altruism, or towards reciprocation-focused “behaviorist helping”. The core driving principle behind these evolutionary pressures is the fact that altruists

---

<sup>11</sup> Someone might be tempted to respond here by arguing either that the desire to reciprocate past help is in fact targeted at particular organisms—i.e. reciprocators—or that the desire to help one's offspring is in fact targeted at particular kinds of behaviors—i.e. offspring-helping behaviors. If so, then these two cases—helping offspring and helping reciprocators—should be seen to be fully parallel. I think this response is coherent, but it also does not affect the main conclusion of my argument: namely, that there are cognitive-efficiency-based reasons to expect some organisms to evolve away from being purely egoistically motivated to help others. Apart from that, though, I also think that it is more straightforward to see the ultimate desires driving altruists and the ones driving reciprocators as importantly different in content (but I am happy to concede that this may be at least partly a matter of convention).

and reciprocation-focused “behaviorist helpers” are taking mental shortcuts relative to egoists: instead of calculating, every time an offspring or a reciprocator is in need of help, whether helping will increase the organism’s own well being, altruists and behaviorist helpers simply presuppose that helping is appropriate. Making this presupposition can be adaptive, for the costs of making this presupposition—which include the (potential) costs of unnecessary helping as well as the absence of any of the benefits that purely egoistic decision making brings (such as the powers of causal reasoning, behavioral flexibility, and memory gains)—can be less than the costs of the constant representational inferences the egoist needs to make—which include losses in decision making speed, and more need for cognitive resources like concentration and attention.<sup>12</sup>

## **V. An Extreme Case: The Adaptiveness of Reflex-Based Helping**

The argument concerning the evolution of altruism and reciprocation-focused “behaviorist” helping can, in some extreme cases, lead to an especially drastic conclusion: namely, that the organism ceases to be motivated by representational, system 2-type considerations altogether, and becomes driven by reflexive, system 1-type responses instead. In turn, this (again) represents an evolution away from egoism, but not towards altruism.

To see this, return to the case of reciprocation-focused “behaviorist helpers”, but also assume that the organism in question reciprocates help in exactly the same way at all times, and that this

---

<sup>12</sup> This point can also be made using Hamilton’s rule. Organisms that make all helping decisions by consulting Hamilton’s rule—which is provably equivalent to making decisions by consulting what maximizes inclusive fitness (Frank, 1998)—will never make the adaptively wrong helping decision, but also face the cost of many relatively complex calculations (the costs and benefits of the helping behavior need to be assessed, and the relatedness to the recipient needs to be estimated). Organisms that instead simply presume that helping offspring or reciprocators is appropriate can, overall, act more adaptively for (i) they have an easier time deciding whether and how to help the organism in need, and (ii) deviations from Hamilton’s rule will be minimal (assuming that helping offspring or reciprocators is indeed sufficiently often adaptive).

is also the only kind of help that is looked for.<sup>13</sup> So, for example, in the much discussed case of food-sharing vampire bats (see e.g. Carter & Wilkinson, 2013), it appears that the only way in which help is demanded or reciprocated is by regurgitating blood.<sup>14</sup> If that is so, though, then it is plausible that these bats do not need to rely on any kind of representational mental state to make a helping decision: they can just react to the perception of their conspecific's need by *doing* the relevant behavior—i.e. by regurgitating blood.

In a bit more detail, the idea here is the following. By the argument of the previous section, there are cases where there is no benefit in making a representational inference to a desire to engage in (reciprocal) helping behavior: after all, the organism always concludes in the same way—namely, that it ought to help. However, if the helping behavior that is offered and asked for is always the same, there is also no value in making a representational inference about how to best provide this help: after all, the organism always concludes in the same way here, too—namely, that it ought to help in the specific way in question. Hence, the organism is structured most efficiently when it avoids any kind of system 2-based reasoning in this context, and just relies on system 1 (see also Millikan, 2002; Dickinson & Balleine, 2000; Schulz, 2011b, 2013). Again, it is important to keep in mind here that it need not be the case that such reflex-driven behavior is literally always adaptive; the point is just that acting reflexively here is overall more adaptive than deliberating about it.

In a nutshell: it is sometimes adaptive to completely avoid using representational mental states to decide whether and how to help others, and to just rely on non-representational reflexes

---

<sup>13</sup> In principle, this point could also arise for offspring-directed (non-reciprocal) help. So for example, for many mammals (including humans), shielding infant offspring with their own bodies from physical dangers is generally adaptive: infant offspring bodies are often much less robust than adult bodies, and can thus suffer great damage from even minor collisions. Since a desire to shield the infant offspring with one's own body is very close to a motor command already, it is thus reasonable to think that reflexively shielding infants in this way is adaptive. However, this case seems less clear cut than the reciprocation-based one in the text, so I focus on the latter one here.

<sup>14</sup> Male vampire bats also seem to reciprocate by grooming (Carter & Wilkinson, 2013). However, this may be a separate case of reciprocation.



instead. Importantly, this (again) amounts to an evolution away from egoism but not towards altruism, for, as noted earlier, egoism and altruism presuppose the presence of ultimate desires, which do not underlie the organism's motivation to help here. Put differently: this is an evolution towards "reflexive helping".

## **VI. Conclusion**

I have tried to show that, when it comes to evolutionary biological accounts of the psychology of helping behavior, it is useful to consider the situation from the point of view of what is most cognitively efficient, and not just of what is most reliable. When doing this, it becomes clear that there is adaptive pressure on at least some organisms to move away from being purely egoistically motivated, and also that this pressure can push in different directions: towards altruism, reciprocation-focused "behaviorist helping", or reflex-driven helping. Note also that it is entirely possible that the same organism is altruistic in some cases (when it comes to offspring), a representational or reflex-driven "behaviorist helper" in other cases (when it comes to reciprocation), and egoistic in yet a third set of cases (when it comes to helping strangers). Equally, it is possible that an organism is just an egoistic helper, or just an altruistic helper, or just a "behaviorist helper". Exactly which of this should be expected to be the case for a particular kind of organism depends on which kinds of decisions (or which parts of its decisions) the organism best makes non-representationally and which it best makes representationally.<sup>15</sup>

Apart from these direct conclusions concerning the evolution of the psychology of helping behavior, I have also suggested that the role evolutionary biological considerations can play in this dispute is an evidential one. While these considerations are unlikely to be able to fully

---

<sup>15</sup> I here do not want to settle the question of how common altruists, behaviorist helpers, and reflexive helpers are, either in relation to each other or in relation to egoists. The point is just that it is plausible that there are some organisms that are altruists, behaviorist helpers, and reflexive helpers. See also Jensen (2012, 316).

confirm some psychological hypotheses, they can lend further support to the latter. In this way, I hope to not only have made a substantive contribution to the debate surrounding the psychology of helping behavior, but also to have added some illumination to some methodological questions about how this debate is best conducted.

## Bibliography

- Batson, Daniel (1991). *The Altruism Question: Toward a Social-Psychological Answer*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brandon, Robert (1990). *Adaptation and Environment*. Princeton: Princeton University Press.
- Brown, Gordon; Wood, Alex, and Chater, Nick (2012). 'Sources of Variation within the Individual'. In P. Hammerstein & J. Stevens (eds.). *Evolution and the Mechanisms of Decision Making*. Cambridge, MA: MIT Press, pp. 227-241.
- Buller, David (2005). *Adapting Minds*. Cambridge, MA: MIT Press.
- Campbell, Richmond, and Kumar, Victor (2012). 'Moral Reasoning on the Ground'. *Ethics* 122: 273-312.
- Carey, Susan, and Spelke, Elizabeth (1996). 'Science and Core Knowledge'. *Philosophy of Science* 63: 515-533.
- Carter, Gerald, and Wilkinson, Gerald (2013): 'Food Sharing in Vampire Bats: Reciprocal Help Predicts Donations More than Relatedness or Harassment'. *Proceedings of the Royal Society B* 280: 20122573.
- Chaiken, Shelly, and Trope, Yaacov (eds.) (1999). *Dual-Process Theories in Social Psychology*. New York: The Guildford Press.
- Chevrand, James, and Wolf, Jason (2009). 'The Genetics and Evolutionary Consequences of Maternal Effects'. In D. Maestriperi & J. Mateo (eds.). *Maternal Effects in Mammals*. Chicago: University of Chicago Press, pp. 11-37.
- Churchland, Patricia (2011). *Braintrust: What Neuroscience Tells Us About Morality*. Princeton: Princeton University Press.

- Clavien, Christine, and Chapuisat, Michel (2013). 'Altruism Across Disciplines: One Word, Multiple Meanings.' *Biology and Philosophy* 28: 125-140.
- Clark, Andy (1997). *Being There*. Cambridge, MA: MIT Press.
- Clutton-Brock, Tim (2009). 'Cooperation Between Non-Kin in Animal Societies'. *Nature* 462: 51-57.
- Curley, James, and Keverne, Eric (2005). 'Genes, Brains and Mammalian Social Bonds'. *Trends in Ecology and Evolution* 20: 561-567.
- Dawkins, Richard (1986). *The Blind Watchmaker*. New York: Norton.
- Dayan, Peter (2012). 'Robust Neural Decision Making'. In P. Hammerstein & J. Stevens (eds.). *Evolution and the Mechanisms of Decision Making*. Cambridge, MA: MIT Press, pp. 151-167.
- Dickinson, Anthony, and Balleine, Bernard (2000). 'Causal Cognition and Goal-Directed Action'. In: C. Heyes and L. Huber (eds.). *The Evolution of Cognition*. Cambridge, MA: MIT Press, pp. 185-204.
- Dickinson, Jeffrey, and Dyer, Fred (1996). 'How Insects Learn About the Sun's Course: Alternative Modeling Approaches'. *Animals to Animals: 4 Proceedings of SAB96*: 193-203.
- Davidic, J., Piliavin, J., Schroeder, D., and Penner, L. (2006). *The Social Psychology of Prosocial Behavior*. Mahawan, N.J.: Lawrence Erlbaum Associates.
- Dretske, Fred (1988). *Explaining Behavior*. Cambridge, MA: MIT Press.
- Epstein, S., Lipson, A., Holstein, C., and Huh, E. (1992). 'Irrational Reactions to Negative Outcomes: Evidence for Two Conceptual Systems'. *Journal of Personality and Social Psychology* 62: 328-339.
- Fehr, Ernst, and Gaechter, Simon (2000). 'Fairness and Retaliation: The Economics of Reciprocity'. *The Journal of Economic Perspectives* 14: 159-181.

- Fodor, Jerry (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Fodor, Jerry (1990). *The Theory of Content*. Cambridge, MA: MIT Press.
- Frank, S. A. (1998). *Foundations of Social Evolution*. Princeton: Princeton University Press.
- Garson, Justin (forthcoming). 'Two Types of Psychological Hedonism'.
- Glimcher, Paul, Dorris, Michael, and Bayer, Hannah (2005). 'Physiological Utility Theory and the Neuroeconomics of Choice'. *Games and Economic Behavior* 52: 213-256.
- Godfrey-Smith, Peter (1996). *Complexity and the Function of Mind in Nature*. Cambridge: Cambridge University Press.
- Goldman, Alvin (1970). *A Theory of Human Action*. Princeton: Princeton University Press.
- Grafen, Alan (1999). 'Formal Darwinism, the Individual-as-Maximizing-Agent Analogy and Bet-Hedging'. *Proc. R. Soc. Lond. B* 266: 799-803.
- Grau, J. (2002). 'Learning and Memory Without a Brain'. In: M. Bekhoff, C. Allen, and G. Burghardt (eds.). *The Cognitive Animal*. Cambridge, MA: MIT Press, pp. 77-88.
- Greene, Joshua (2008). 'The Secret Joke of Kant's Soul'. In W. Sinnott-Armstrong (ed.). *Moral Psychology*. Vol. 3. Cambridge, MA: MIT Press, pp. 35-79.
- Haidt, Jonathan (2001). 'The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment'. *Psychological Review* 108: 814-834.
- Haldane, John (1932). *The Causes of Evolution*. Princeton: Princeton University Press.
- Hammerstein, Peter (2003). 'Why is Reciprocity so Rare in Social Animals? A Protestant Appeal.' In P. Hammerstein (ed.). *Genetic and Cultural Evolution of Cooperation*, Cambridge, MA: MIT Press, pp. 83-93.
- Hausfater, Glenn, and Hrdy, Sarah Blaffer (eds.) (1984). *Infanticide: Comparative and Evolutionary Perspectives*. Chicago: Aldine Transactions.

- Hrdy, Sarah Blaffer (1999). *Mother Nature*. New York: Ballentine.
- Jensen, Keith (2012). 'Who Cares? Other-Regarding Concerns – Decisions with Feeling'. In P. Hammerstein and J. Stevens (eds.). *Evolution and the Mechanisms of Decision Making*. Cambridge, MA: MIT Press, pp. 299-317.
- Kacelnik, Alex (2012). 'Putting Mechanisms into Behavioral Ecology'. In P. Hammerstein & J. Stevens (eds.). *Evolution and the Mechanisms of Decision Making*. Cambridge, MA: MIT Press, pp. 21-38.
- Kahneman, Daniel (2003). 'Maps of Bounded Rationality: Psychology for Behavioral Economics.' *American Economic Review* 93: 1449-1475.
- Kitcher, Philip (2011). *The Ethical Project*. Cambridge, MA: Harvard University Press.
- Kruglanski, Arie, and Gigerenzer, Gerd (2011). 'Intuitive and Deliberate Judgments Are Based on Common Principles'. *Psychological Review* 118: 97-109.
- Lewontin, Richard (1998). 'The Evolution of Cognition: Questions We Will Never Answer.' In D. Scarborough and S. Sternberg (eds.). *Methods, Models, and Conceptual Issues: An Invitation to Cognitive Science*, Vol. 4. Cambridge, MA: MIT Press, pp. 107-132.
- Machery, Edouard (forthcoming). 'Discovery and Confirmation in Evolutionary Psychology'. In J. Prinz (ed.). *Oxford Handbook of Philosophy of Psychology*. Oxford: Oxford University Press.
- Mackintosh, N. J. (1994). *Animal Learning and Conditioning*. San Diego: Academic Press.
- McFarland, D. (1996). 'Animals as Cost-Based Robots'. In M. Boden (ed.). *The Philosophy of Artificial Life*. Oxford: Oxford University Press, pp. 179-207.
- Millikan, Ruth (1984). *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.

- Millikan, Ruth (2002). *Varieties of Meaning*. Cambridge, MA: MIT Press.
- Mithen, Steven (1990). *Thoughtful Foragers: A Study of Prehistoric Decision Makers*.  
Cambridge: Cambridge University Press.
- Nagel, Thomas (1970). *The Possibility of Altruism*. Princeton: Princeton University Press.
- Nettle, Daniel (2012). 'Error Management. In P. Hammerstein & J. Stevens (eds.). *Evolution and the Mechanisms of Decision Making*. Cambridge, MA: MIT Press, pp. 69-80.
- Okasha, Samir (2006). *Evolution and the Levels of Selection*. Oxford: Oxford University Press.
- Papineau, David (1987). *Reality and Representation*. Oxford: Blackwell.
- Papineau, David (2003). *The Roots of Reason*. Oxford: Oxford University Press.
- Penke, L. (2010). 'Bridging the Gap Between Modern Evolutionary Psychology and the Study of Individual Differences'. In D. M. Buss & P. H. Hawley (Eds.). *The Evolution of Personality and Individual Differences*. Oxford: Oxford University Press, pp. 243-279.
- Prinz, Jesse (2002). *Furnishing the Mind: Concepts and their Perceptual Basis*. Cambridge, MA: MIT Press.
- Richardson, Robert (2007). *Evolutionary Psychology as Maladapted Psychology*. Cambridge, MA: MIT Press.
- Sachs, Joel; Mueller, Ulrich; Wilcox, Thomas, and Bull, James (2004). 'The Evolution of Cooperation'. *The Quarterly Review of Biology* 79: 135-160.
- Schulz, Armin (2011a). 'Sober & Wilson's Evolutionary Arguments for Psychological Altruism: A Reassessment'. *Biology and Philosophy* 26: 251-260.
- Schulz, Armin (2011b). 'The Adaptive Importance of Cognitive Efficiency: An Alternative Theory of Why We Have Beliefs and Desires.' *Biology and Philosophy* 26: 31-50.

- Schulz, Armin (2012). 'Heuristic Evolutionary Psychology'. In K. Plaisance & T. Reydon (eds.). *Philosophy of Behavioral Biology*. Dordrecht: Springer, pp. 217-234.
- Schulz, Armin (2013). 'The Benefits of Rule Following: A New Account of the Evolution of Desires'. *Studies in History and Philosophy of Biological and Biomedical Sciences* 44: 595-603.
- Slooman, S. A. (1996). 'The Empirical Case for Two Systems of Reasoning'. *Psychological Bulletin* 119: 3-22.
- Sober, Elliott (2001). 'The Two Faces of Fitness'. in R. Singh, D. Paul, C. Krimbas, and J. Beatty (eds.), *Thinking about Evolution: Historical, Philosophical, and Political Perspectives*. Cambridge: Cambridge University Press, pp. 309-321.
- Sober, Elliott (2008). *Evidence and Evolution*. Cambridge: Cambridge University Press.
- Sober, Elliott, and Wilson, David Sloan (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
- Stanovich, K. E., and West, R. F. (2000). 'Individual Differences in Reasoning: Implications for the Rationality Debate.' *Behavioral & Brain Sciences* 23: 645-665.
- Sterelny, Kim (2001). *The Evolution of Agency and Other Essays*. Cambridge, MA: Cambridge University Press.
- Sterelny, Kim (2003). *Thought in a Hostile World*. Oxford: Blackwell.
- Sterelny, Kim (2012). *The Evolved Apprentice*. Cambridge, MA: MIT Press.
- Stich, Stephen (2007). 'Evolution, Altruism and Cognitive Architecture: A Critique of Sober and Wilson's Argument for Psychological Altruism'. *Biology and Philosophy* 22: 267-281.
- Stich, Stephen (forthcoming). 'Why There Might Not Be an Evolutionary Explanation for Psychological Altruism'.



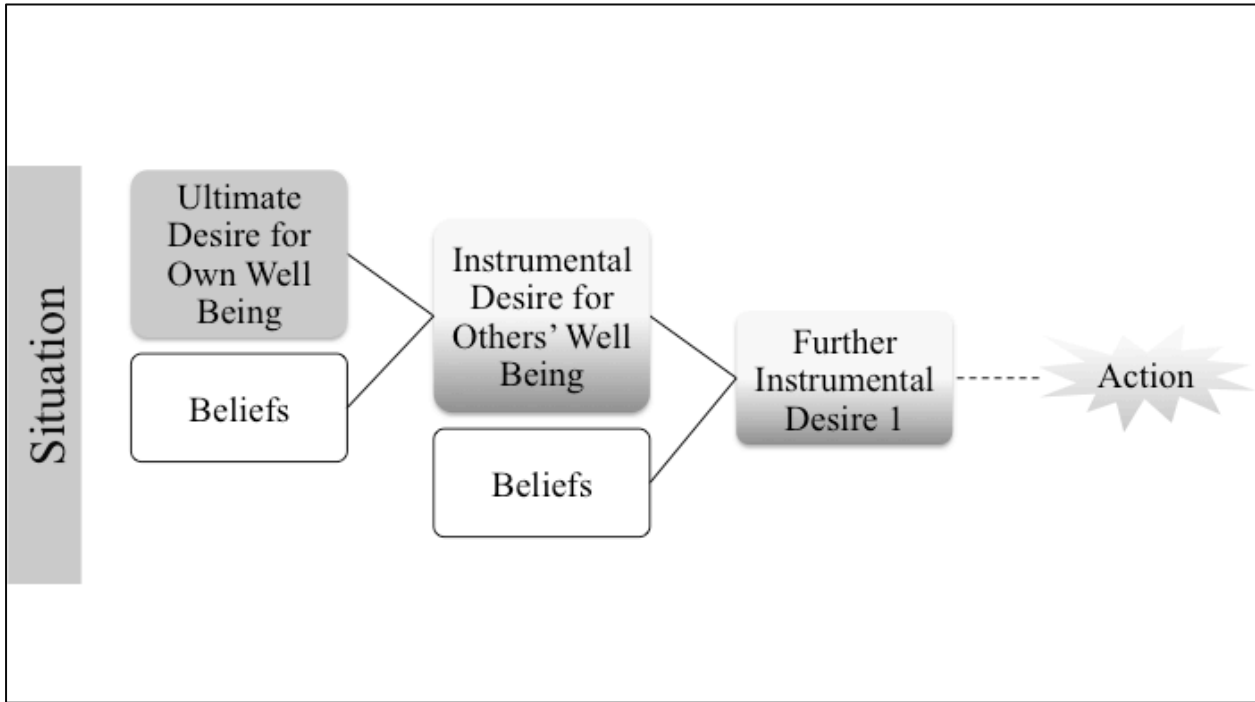
Stich, Stephen, Doris, John, and Roedder, Erica (2010). 'Altruism'. In J. Doris and the Moral Psychology Research Group (eds.). *The Moral Psychology Handbook*. Oxford: Oxford University Press, pp. 147-205.

Thometz, N.; Tinker, M., Staedler, M.; Mayer, K., and Williams, T. (2014). 'Energetic Demands of Immature Sea Otters from Birth to Weaning: Implications for Maternal Costs, Reproductive Behavior and Population-Level Trends'. *The Journal of Experimental Biology* **217**: 2053-2061.

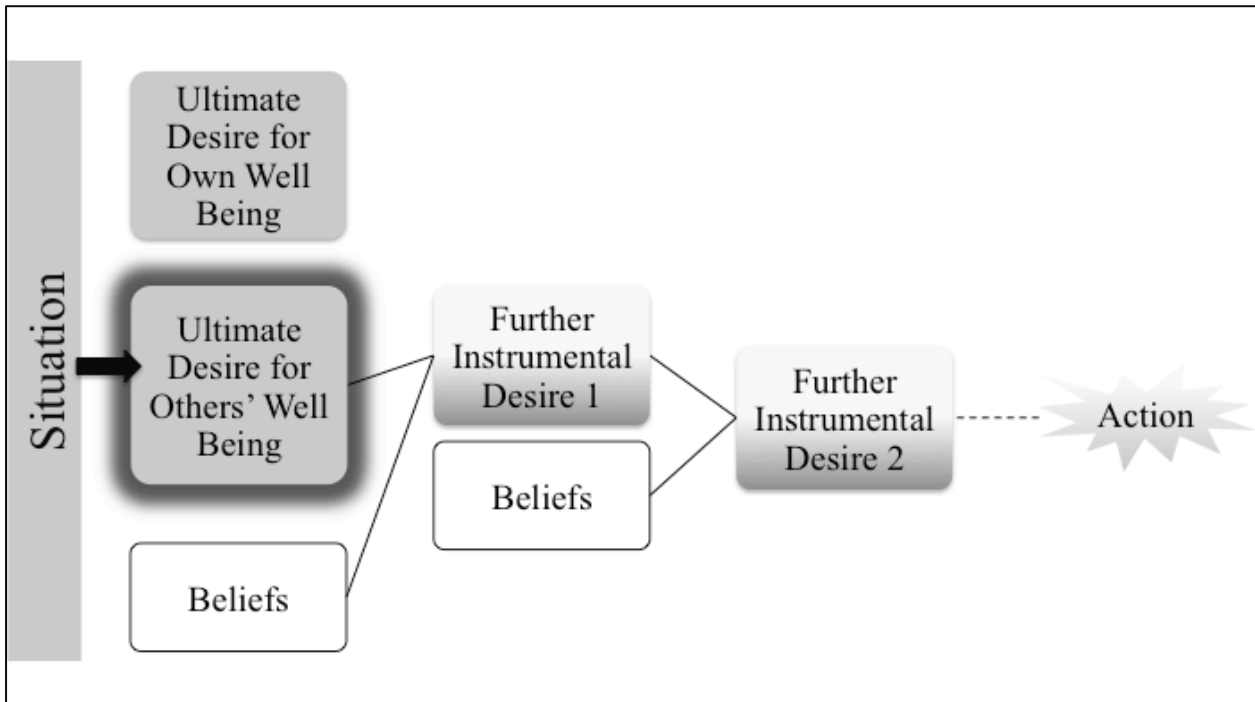
Trivers, Robert (1971). 'The Evolution of Reciprocal Altruism'. *The Quarterly Review of Biology* 46: 35-57.

Trivers, Robert (1974). 'Parent-Offspring Conflict'. *American Zoologist* 14: 247-262.

Figures



[Figure 1: An Egoistic Decision to Help]



[Figure 2: An Altruistic Decision to Help]