Biol Philos
DOI 10.1007/s10539-009-9179-5

# Sober & Wilson's evolutionary arguments for psychological altruism: a reassessment

**Armin W. Schulz**

Received: 2 May 2009 / Accepted: 18 August 2009


**Abstract** In their book *Unto Others*, Sober and Wilson argue that various evolutionary considerations (based on the logic of natural selection) lend support to the truth of psychological altruism. However, recently, Stephen Stich has raised a number of challenges to their reasoning: in particular, he claims that three out of the four evolutionary arguments they give are internally unconvincing, and that the one that is initially plausible fails to take into account recent findings from cognitive science and thus leaves open a number of egoistic responses. These challenges make it necessary to reassess the plausibility of Sober & Wilson's evolutionary account—which is what I aim to do in this paper. In particular, I try to show that, as a matter of fact, Sober & Wilson's case remains compelling, as some of Stich's concerns rest on a confusion, and those that do not are not sufficiently strong to establish all the conclusions he is after. The upshot is that no reason has been given to abandon the view that evolutionary theory has advanced the debate surrounding psychological altruism.

**Keywords** Altruism · Evolution · Adaptive · Evidence · Sub-doxastic · *Unto Others*

## Introduction

Some ten years ago, Sober & Wilson gave a new impetus to the old debate about whether we are psychological altruists: they argued that *natural selection* favours psychological altruism over psychological egoism, and that this therefore makes it more plausible that we, in fact, are altruistically motivated (at least sometimes).

A. W. Schulz (✉)
Department of Philosophy, University of Wisconsin-Madison,
5185 Helen C. White Hall, 600 N. Park St, Madison, WI 53706, USA
e-mail: awschulz@wisc.edu

Published online: 30 August 2009

Recently, though, their evolutionary approach to this issue has come under fire—Stephen Stich (2007), in particular, has raised some deep challenges to its overall cogency. This makes it necessary to reassess Sober and Wilson's (1998) approach, and to determine whether it can still be used to make the case for psychological altruism. This is what I aim to do in this paper.

To do this, I begin by laying out the presuppositions of Sober & Wilson's arguments in the next section. After that, I present and discuss Sober & Wilson's evolutionary arguments, together with the objections raised by Stich (2007). I conclude in the final section.

## Presuppositions of Sober & Wilson's evolutionary arguments for psychological altruism

In *Unto Others*, Sober & Wilson claim that various evolutionary arguments support the truth of the thesis that we are psychological altruists (which they see as insufficiently well confirmed by purely philosophical or purely psychological considerations—Sober and Wilson 1998, chaps. 8, 9; see also Stich et al. forthcoming). Before it is possible to assess whether they are right in claiming this, though, it is necessary to address two preliminary issues.

The first of these issues concerns the question as to what, exactly, the debate surrounding psychological altruism is about. In order to determine this, it is best to begin by defining the notion of 'psychological altruism' more precisely; this can be done as follows (Sober and Wilson 1998, p. 201)[1]:

*Psychological altruism*: The existence of ultimate desires concerning the well-being of others.

Three aspects of this thesis should be made explicit here.

Firstly, it is *pluralist* in structure: it does not say that *all* of one's ultimate desires are for the well-being of others (this would obviously be false). It merely claims that *some* of one's ultimate desires are for the well-being of others. By contrast, *psychological egoism* (the major alternative thesis it contrasts with) is *monistic* in structure: it claims that *all* of one's ultimate desires are for one's own well-being only (Sober and Wilson 1998, p. 228).

Secondly, the above characterisation of psychological altruism rests crucially on the concept of an 'ultimate' desire. Ultimate desires are opposed to instrumentalist desires, in that the latter are *derived from* the former by means of intervening beliefs or belief-like states. Spelling out what this 'derives from' means is quite tricky, but for present purposes, it is enough to note that it is widely accepted that a 'rough and ready' criterion for a desire to be instrumental is that the *only* reason why the desire is held is that the agent has some particular belief or belief-like state. While this criterion is unlikely to do justice to all cases, it is all that is needed for the discussion

[1] It is important not to confuse psychological altruism with the very different notion of *evolutionary altruism* (the existence of phenotypes that reduce the fitness of an organism, relative to that of the other members of its group). For more on the latter, see the first five chapters of Sober and Wilson (1998).

here (for more on this, see Stich 2007; Sober and Wilson 1998, pp. 217–222; Goldman 1970).

Thirdly, among the egoistic alternatives to altruism, *hedonism* is the most important (see also Sober and Wilson 1998, pp. 297 and 318). Hedonism is the thesis that all of a person's ultimate desires concern *her own pleasure and pain only*. For what follows, it is further useful to distinguish among two particular kinds of hedonism (see also Stich 2007, pp. 273–274): 'Current Pain Hedonism' (CPH) and 'Future Pain Hedonism' (FPH). CPH is directed at the avoidance of *current* pain: an agent's beliefs are assumed to (sometimes) generate pain; the agent is then motivated to act in such a way as to make this pain cease. FPH, by contrast, is directed at the avoidance of *future* pain: the agent is said to have certain beliefs about what *will* cause her pain; she is then motivated to act in a way that prevents this pain from actually coming about.[2]

In this way, the debate surrounding psychological altruism can be reduced to the question as to which of CPH, FPH, and altruism describes our motivational architecture best: if one of the first two turns out to be true, egoism is vindicated; otherwise it is falsified. With this in mind, consider now the second preliminary issue to be addressed here.

This issue concerns the question as to which conclusion, exactly, Sober & Wilson's evolutionary arguments (jointly) ought to be taken to try to establish. In principle, there are three options here:

(A) We ought to *accept* the thesis that we are psychologically altruistic (i.e. the evolutionary considerations entail that the probability of the thesis is *very high*).
(E) There is *evidence for* the thesis that we are psychologically altruistic (i.e. the evolutionary considerations *raise* the probability of the thesis).
(M) Evolutionary considerations are *methodologically helpful* for the evaluation of the thesis that we are psychologically altruistic (i.e. the evolutionary considerations *lead us to new ways of assessing* the probability of the thesis).

Now, a brief look at Sober and Wilson (1998) is enough to make clear that option (A) can be immediately discarded: Sober & Wilson quite clearly do not pretend that their arguments could achieve anything so sweeping as the *establishment of the truth* of the thesis of psychological altruism. Equally clearly, though, conclusion (E) *is* advanced by them: they are explicit in noting that they seek to show that evolutionary theory can provide *considerations speaking in favour of* the truth of altruism—which is all that giving evidence amounts to here (Sober and Wilson 1998, p. 12). The situation with respect to (M) is more ambiguous, however: while it is nowhere *explicitly* endorsed, a number of Sober & Wilson's remarks at least *suggest* acceptance of it (see e.g. Sober and Wilson 1998, pp. 3, 8, 334). For these reasons, it seems best to proceed on the assumption that Sober & Wilson accept *both* (E) and (M).

[2] Note that Sober & Wilson claim that FPH can be ruled out for philosophical reasons alone (Sober and Wilson 1998, pp. 281–287). For present purposes, though, it is better not to follow them in this, and to assume that FPH is still a live option—this is more in line with Stich (2007), and makes the exposition of the arguments below easier.

Given this, what needs to be done now is to see to what extent recent criticisms have thrown into doubt the ability of Sober & Wilson's evolutionary arguments to support these two conclusions. The next section aims to do this.

## Sober & Wilson's evolutionary arguments for psychological altruism: objections and replies

Sober & Wilson put forward four arguments that suggest that altruism is a more *reliable*—and thus, a more *adaptive*—motivational system than hedonism for causing parents to help their children (Sober and Wilson 1998, pp. 305–306).[3] However, in the last five years, a number of criticisms have been raised against these arguments—in particular, Stephen Stich (2007) has mounted a powerful and insightful attack on their cogency. While he is not alone in criticising Sober & Wilson's arguments (see e.g. Lemos 2004), his objections are the most clear and cogent ones voiced thus far; because of this, they are taken to be the main target in the discussion to follow.

In this, though, it needs to be noted that not *all* of these objections need to be rebutted for Sober & Wilson's case against hedonism to remain compelling. Since Sober & Wilson present four arguments to this effect, their *overall* defence of altruism will be vindicated as long as *one* of these arguments remains compelling. That said, it also needs to be noted that, since there are *two* conclusions that Sober & Wilson are assumed to be aiming at—(E) *and* (M)—it is sufficient for Stich to rule out *one* of these for all four of their arguments. With this in mind, I now present these four arguments, followed by Stich's objections to them; in each case, I then consider various replies that could be given to these objections.[4]

### Argument 1: the belief/emotion link

Sober and Wilson (1998) note that hedonistic organisms depend on the fact that, in situations where helping behaviour is adaptive (e.g. when it comes to one's own children), emotions like distress or pain are generated. The problem with this, according to Sober & Wilson, is that, at times, these emotions might *not* be generated: no emotion-generator is completely fail-save, so that there is always the possibility that, even though the agent believes that her child needs help, she does not feel pain or distress. Since helping the child is (by definition) adaptive, the hedonist will therefore, at times, be led to forgo possible adaptive benefits.

[3] A brief remark about why Sober & Wilson focus on *reliability* and *parental care*. They focus on *reliability*, as they think that considerations of availability and energetic efficiency do not distinguish between the two motivational architectures (Sober and Wilson 1998, pp. 221–223). While Lemos (2004) calls this into question, it is best to grant this assumption here. They focus on *parental care*, as they (reasonably) think that the latter is likely to be adaptively important *to the parent*: helping one's children, by and large, will increase one's *own* fitness as well as that of one's children (see also Stich 2007, p. 270).

[4] I follow Stich (2007) in the way the arguments are presented, ordered and numbered (which differs slightly from how they are laid out in Sober and Wilson 1998). Doing this should not introduce any infelicities—in particular, there is no reason to think that Stich (2007) has misread Sober and Wilson (1998) in any way.

Importantly, though, the altruist is immune to this problem: she has an ultimate desire to help, and thus does not require the generation of the appropriate emotions. Hence, the altruist does not depend on this generation working reliably—and will thus be fitter than the hedonist (Sober and Wilson 1998, pp. 315–316).

For two reasons, though, Stich finds this argument to be unconvincing. Firstly, he notes that there are many cases where less than fully reliable motivational systems have evolved (Stich 2007, pp. 275–276). For example, it seems it would be highly adaptive if, instead of having to rely on a link between *believing ourselves to be danger* and *being afraid* to avoid dangerous situations, we could just rely on innate dispositions to do so; evidently, however, this is not what has evolved. Secondly, Stich notes that FPH is immune from Sober & Wilson's reasoning, since it requires no belief/emotion link either. As noted earlier, FPH is based on the idea that certain situations induce beliefs about what the hedonist *will feel in the future*—however, the emotion itself does not actually need to be *generated*. For this reason, this form of hedonism remains tenable, even if the above argument is accepted (Stich 2007, p. 275 note 7).

Upon closer consideration, though, neither part of Stich's criticism can stand up to scrutiny. The first part fails, as it rests on a conflation of what is *adaptive* with what *evolves*. To see this, remember that Sober & Wilson's aim in appealing to natural selection is—at most—only to present considerations that *favour* the evolution of altruism over that of hedonism. Because of this, it is irrelevant to point out (as Stich has done) that in many cases less than fully reliable systems *evolved*: what matters is only which kind of system is more *adaptive*. Unless we have reason to think that the more adaptive system *never* evolves (or at least that it did not evolve *in the present case*), Stich's criticism thus does not impinge on Sober & Wilson's case for psychological altruism at all. Importantly, it furthermore needs to be noted that Stich has done nothing to shed doubt on the evidential importance of what is adaptive for what can be expected to evolve (either in general or when it comes to altruism): he merely states that *sometimes*, less than fully reliable systems have evolved. This, though, is entirely consistent with it being the case that, in general, the reliability of a system *is* evidence for its evolution. Because of this, Stich's first worry does not show that Argument 1 fails to present evidence against CPH.

The situation with respect to the second part of his criticism is more ambiguous, however. On the one hand, it must clearly be acknowledged that Stich is right in noting that Argument 1 leaves FPH untouched. For this reason, the support this argument confers to the altruism thesis can only be *weak*: it can only be seen to favour altruism to the extent that the latter is contrasted with CPH. On the other hand, though, it also needs to be acknowledged that Argument 1 still makes clear that, *if* hedonism is to be evolutionarily plausible, it has to be of the FPH variety. This is important, as it means that this argument succeeds in constraining the space of possible hedonistic architectures, and thus makes testing the truth of altruism easier. In particular, the argument at least shows that, in order to see whether we are altruists, we need to test which ultimate desires *about the future* motivate our actions. This is helpful, as this future-directedness of the issue (a) was not clear beforehand (see e.g. Batson 1991), and (b) significantly limits the sprawling set of hedonistic hypotheses that need to be considered to empirically investigate whether

we are altruists (see e.g. Stich et al. forthcoming, pp. 66–67). In this way, Argument 1 can *lead to* the further confirmation of psychological altruism, and thus advances the debate surrounding this thesis.

Argument 2: the need for producing sufficient amounts of the relevant emotions

Following on from Argument 1, Sober & Wilson note that hedonism also requires that *sufficient amounts* of the relevant motivating emotions (pain or distress) are produced for the appropriate kind of helping behaviour to come about. However, given the fact that these necessary amounts of pain and distress (plausibly) differ from situation to situation, generating them requires much 'tricky engineering'. This matters, as the altruistic alternative does *not* require this tricky engineering—and will therefore be simpler for natural selection to build. In turn, this makes it more adaptive than the hedonist alternative (Sober and Wilson 1998, pp. 315–316).

There are three reasons why Stich thinks Sober &Wilson's second argument is not very compelling. Firstly, there are many cases where the relevant kind of 'tricky engineering' between cause and effect has been achieved in evolution; there is thus no reason to think that it could not have been done here (Stich 2007, p. 276). Secondly, this argument again does not apply to FPH (Stich 2007, p. 275 note 7). Thirdly, the same tricky engineering is needed on Sober & Wilson's own account as well: after all, for altruism to be a more reliable motivator, the ultimate desire for the welfare of the child must also be ensured to be stronger than other ultimate desires—like maximising one's own pleasure. This means that this is not an argument that can distinguish between hedonism (in either form) and altruism (Stich 2007, p. 276 note 8).

Overall, Stich is right in dismissing this argument as being unhelpful to Sober & Wilson's case. The reason for this rests primarily on the cogency of Stich's third worry: the kinds of considerations put forward by Sober & Wilson just do not distinguish altruism from any of the forms of hedonism under consideration. Hence, this argument will not be further discussed here.

Argument 3: pluralism versus monism

Thirdly, Sober & Wilson return to the fact that hedonism is a monist system, whereas altruism is a pluralist one. This matters, as Sober & Wilson then argue that, ceteris paribus, pluralist systems are more reliable than monist ones: they can function even if they are partly damaged. In particular, in the present context, an altruist might engage in adaptive helping behaviour even if her ultimate desire to do so is, for some reason, ineffectual in bringing this about (e.g. she might also be moved to help by the hedonist desire to avoid feeling guilty later on). Accordingly, Sober & Wilson conclude that, ceteris paribus, natural selection is likely to favour altruism over hedonism (Sober and Wilson 1998, pp. 320–321).

Stich's concern with Sober & Wilson's third argument can be stated very briefly: there are lots of cases where essential systems have evolved that are *not* pluralist (e.g. the heart and the liver). Hence, he concludes, the same might be true of altruism (Stich 2007, pp. 276–277).

This objection, though, is not very compelling—and that for the same reasons that Stich's attack on Argument 1 was unconvincing: it falls prey to the conflation between evolution and adaptiveness. However, it is worthwhile to dig a little deeper here—for this objection runs into difficulties *even if* it is reformulated so that it avoids collapsing the evolution/adaptation distinction. To see this, note that this reformulation can be done in two different ways.

In the first reformulation, Stich accepts Sober & Wilson's contention that pluralism is more adaptive than monism *ceteris paribus*, but claims at the same time that there are reasons for thinking that, *in the case at hand* (i.e. that of psychological altruism), the pluralist system was not more adaptive *overall*. However, for one simple reason, this reconstruction is unlikely to save Stich's objection from refutation: it is not backed up by sufficient evidence.

This comes out clearly from noting that Stich does not, in fact, present any particular considerations that suggest that other factors (be they selectionist or non-selectionist) outweighed the adaptive value of the altruistic-pluralistic organisation. All he does is note that there may be *costs* that come from building a pluralist system (Stich 2007, p. 277); however, this sort of general sceptical challenge is unconvincing unless it is made clearer where these costs come from and why they are likely to be greater than the benefits of a pluralist design. This is particularly important here, since there is no obvious candidate for what these costs might be—not all pluralist systems contain more functional or material elements than alternative monist systems, and neither do they *necessarily* use up more energy. Because of this, Stich's appeal to the mere *possibility* of pluralist systems being costly does not challenge the cogency of Sober & Wilson's third argument.

In the second reformulation of Stich's objection, he simply *denies* the claim that pluralist systems are more adaptive than monist ones in the environments humans evolved in (even ceteris paribus): just as being tall may not have been more adaptive than being short, there was no major adaptive benefit to the organism's having a pluralist motivational architecture—this architectural design was selectively neutral. However, this reformulation, too, cannot save Stich's objection—and this is again because there is a lack of evidence to contend with.

In particular, while it *may* be true that, in the relevant environments, pluralism does not have a higher adaptive value than monism, Stich has done nothing to establish this. Just claiming that pluralism *may not* be more adaptive than monism (in the relevant environments) is not very compelling—especially since Sober & Wilson have motivated their position with further arguments (e.g. the fact that pluralism leads to increased *reliability* in adaptive action generation). At least as matters stand currently, therefore, there are good reasons to think that pluralist systems *are* (ceteris paribus) adaptive—and none to think that they are not.

At this point, Stich might offer the following counter-reply.[5] Even if it is granted that Argument 3 provides evidence for altruism over hedonism, this evidence seems *extraordinarily weak*. This can be seen from the fact that relying on the reasoning embodied in this argument leads one to make many false inferences: for example, it seems to *falsely* suggest that we have two hearts—after all, the heart is an essential

[5] In fact, this counter-reply *was* offered to me by Stephen Stich (in personal communication).

part of the human organism, and we would seem to profit greatly from having a 'backup blood pump'. This sort of mistaken prediction makes vivid the fact that adaptationist arguments of the above form should not be given much credence in our inferences about what has actually evolved: *at best*, they make for extremely uninformative evidence.

This counter-reply, though, fails to be convincing as well. To see this, note that the reason why we have evolved just one heart—granting, for the sake of the argument, that having two would have been more adaptive—may merely be that natural selection is not the only relevant factor for the evolution of the heart. This organ has been part of the mammalian branch for a long time, and *inheritance* may be a major part of the explanation for its non-pluralist design. Importantly moreover, the *importance* of natural selection relative to other evolutionary factors can be assessed separately from the *target* of the relevant selective regime. To be sure, finding out about what is adaptive will *gain* in importance the more powerful we think natural selection was in influencing the evolution of the trait in question; however, this does not mean that, initially, this importance is nil. In particular, given the ubiquity of natural selection in evolution (Orzack and Sober 1994), uncertainty about the strength of the relevant selective pressures does not entail the irrelevance of adaptationist arguments.

Argument 4: the possibility of maladaptive updating

Sober & Wilson's final argument begins by noting that for hedonism to motivate helping behaviour, it is necessary that the hedonist believes that helping her child is the best way to alleviate her current distress. However, such beliefs—like all beliefs—can be removed from the cognitive system through the receipt of further evidence. For example, when realising that taking drugs is a quicker and more immediately rewarding way to relieve her discomfort at believing her child to be in danger, the hedonist may be led to cease to engage in the helping behaviour. Altruism, by contrast, cannot fall prey to this kind of untoward updating—it relies on ultimate desires to generate the motivation to help, and is therefore immune to these worries. Accordingly, altruism is the more reliable, and therefore the more adaptive, cognitive architecture (Sober and Wilson 1998, pp. 314–315).

Stich's worries surrounding this argument centre on the fact it does not take into account certain recent advances in cognitive science (Stich 2007, p. 267). In particular, he notes that much of recent cognitive science appeals to so-called 'sticky states' (or, in his earlier terminology, 'sub-doxastic states'—Stich 1978). Sub-doxastic states are mental states that function in many ways like beliefs (e.g. they represent what the state of the world is), but which are otherwise quite different from the rest of the agent's beliefs. In particular, they are not introspectible, often fail to be inferentially integrated, and—for present purposes most importantly—cannot be removed easily from the agent's cognitive system (Stich 2007, pp. 278–279). In other words, while sub-doxastic states often play the same causal roles that beliefs do, they differ from the latter in that they remain part of our cognitive system *even if* evidence of their falsity is received.

The importance of these sticky states for the current discussion is that they give defenders of hedonism the option to respond to Sober & Wilson's fourth argument by claiming that the mental states that ground the desire for helping behaviour are, in fact, sub-doxastic states. For this reason, they will be just as hard to remove from the cognitive system as ultimate desires to help are—thus denying Sober & Wilson the key premiss in their argument (Stich 2007, p. 280).

There are two quick possible responses to Stich's objection that are not very compelling, but which it is useful to mention anyway, as they clarify the nature of his objection. Firstly, questioning the *existence* of sticky states is implausible. While it may be true that the views of those defending their existence—like Chomsky, Carey, Spelke, and Fodor—have been criticised, they do command wide acceptance, and denying their cogency outright is not a compelling option at this point. Secondly, neither is it very convincing to claim that, since sticky states are immune to evidential updating, the desires they give rise to must be *ultimate* ones (and thus that the above is not an argument against altruism). To see this, note that these desires are in the agent's cognitive system *only* due to the fact that she also has certain belief-like states (which just happen to be sub-doxastic ones). This is crucial, as it means that the desires satisfy the above criterion of instrumentality: the organism would not have them, if the belief-like states they are based on would somehow be removed from her cognitive system. Furthermore, there is no reason to think that this assessment would change if one were to make use of a more sophisticated criterion of instrumentality (see also Stich 2007, p. 279).

Fortunately, Sober & Wilson have a much more plausible response to Stich's objection available to them: namely, to simply *accept* it. They can do this, as it now greatly matters which of their two conclusions—(E) or (M)—Stich is after: in particular, while Stich may have managed to show that Argument 4 does not present *evidence* in favour of altruism, he has not succeeded in showing that the argument cannot play *any role* in the debate surrounding altruism *at all*. In fact, by bringing into view the fact that, for hedonism to be evolutionarily plausible, it must rest on sub-doxastic states, the debate about altruism is pushed forward considerably: it is now possible to test whether people's seemingly altruistic behaviours rest on belief-like mental states that are non-introspectible, inferentially isolated, and not sensitive to evidence.

This marks a significant step forward, as several well-known methods have been devised to test for the presence of exactly these kinds of states. For example, in developmental psychology, the 'differential looking times paradigm' has been used to show that human infants seem to be born with various sub-doxastic belief-like states (see e.g. Carey 1998). On top of this, cognitive psychologists have constructed various tests for the presence of non-introspectible beliefs in adults (see e.g. Pylyshyn 1999). All of these methods are now shown to be relevant to the debate surrounding altruism—something that was not clear beforehand, and which marks a significant step forward in this debate.

Where does all of this leave Stich's attack on Sober & Wilson's evolutionary arguments? After the dust has settled, the dialectical situation turns out to be as follows:

(1) In the context of Argument 1, Stich has partially falsified (E) (when altruism is contrasted with FPH), but he has failed to falsify (M).
(2) In the context of Argument 2, he has falsified both (E) and (M).
(3) In the context of Argument 3, he has failed to falsify (E).
(4) In the context of Argument 4, he has falsified (E), but he has failed to falsify (M).

Now, as noted earlier, in order to be successful, Stich's objections needed to establish that at least one of (E) and (M) does not follow from *any* of Sober & Wilson's arguments. Since they have not achieved this, though, Stich has given no reason to think that Sober & Wilson's evolutionary arguments do not advance the debate surrounding psychological altruism—either evidentially (due to Argument 3 and partially Argument 1) or methodologically (due to Arguments 1 and 4).

## Conclusion

I have tried to show that Sober & Wilson's evolutionary arguments for psychological altruism remain compelling. In particular, I have tried to make clear that, while raising several important issues, Stich's (2007) objections fail to show that Sober & Wilson's evolutionary arguments are not evidentially or methodologically useful to the evaluation of the truth of psychological altruism. Importantly, I thus hope to also have shown more generally when and why evolutionary reasoning has much value in psychology and philosophy.

**Acknowledgments** I thank Elliott Sober, Stephen Stich, Kim Sterelny, and an anonymous referee of this journal for useful comments on earlier drafts of this paper.

## References

Batson CD (1991) The altruism question: towards a social-psychological answer. Lawrence Erlbaum Associates, Hillsdale
Carey S (1998) Knowledge of number: its evolution and ontogenesis. Science 242:641–642
Goldman A (1970) A theory of human action. Prentice-Hall, Princeton
Lemos J (2004) Psychological hedonism, evolutionary biology, and the experience machine. Philos Soc Sci 34:506–526
Orzack S, Sober E (1994) Optimality models and the test of adaptationism. Am Nat 143:361–380
Pylyshyn Z (1999) Is vision continuous with cognition? Behav Brain Sci 22:341–423
Sober E, Wilson DS (1998) Unto others: the evolution and psychology of unselfish behavior. Harvard University Press, Cambridge, MA
Stich S (1978) Beliefs and sub-doxastic states. Philos Sci 45:499–518
Stich S (2007) Evolution, altruism and cognitive architecture: a critique of Sober and Wilson's Argument for psychological altruism. Biol Philos 22:267–281
Stich S, Doris J, Roedder E (forthcoming) The science of altruism. In: Doris JM, Harman G, Nichols S, Prinz J, Sinnott-Armstrong W, Stich S (eds) The Oxford handbook of moral psychology. Oxford University Press, Oxford