

The Evolution of Psychological Altruism

Gualtiero Piccinini, 1 University Blvd, Department of Philosophy, University of Missouri – St. Louis, St. Louis, MO 6121, piccininig@umsl.edu, <http://www.umsl.edu/~piccininig/>

Armin Schulz, 3101 Wescoe Hall, Department of Philosophy, University of Kansas, awschulz@ku.edu, <http://people.ku.edu/~a382s825/>

Abstract. We argue that there are two importantly different kinds of altruistic motivations: *classical psychological altruism*, which generates ultimate desires to help other organisms at least partly for those organisms' sake, and *nonclassical psychological altruism*, which generates ultimate desires to help other organisms for the sake of the organism providing the help. We then argue that classical psychological altruism is adaptive if the desire to help others is intergenerationally reliable, and thus need not be learned. Nonclassical psychological altruism is adaptive when the desire to help others is adaptively learnable. Thus, both kinds of psychological altruism are likely to be selected for. This theory of the motivational structures underlying helping behaviors opens up new avenues for the productive, interdisciplinary study of psychological altruism.

Some organisms *behave* altruistically—they reduce their own reproductive (direct) fitness (i.e., the expected number of their offspring) while increasing the reproductive (direct) fitness of another organism (West et al., 2007, Okasha, 2013). Such altruistic behavior can be selected for (Sober & Wilson, 1998; Gardner & West, 2010; Okasha, 2006): its inclusive fitness can be positive (Okasha, 2011; Grafen, 2006). (An organism's inclusive fitness is defined as the sum of an organism's direct fitness—i.e., the expected number of its own offspring—and its relatedness-weighted contribution to the direct fitness of every other organism in the population; see West et al., 2007, 2011, Okasha, 2011.)

What is controversial is whether behavior is ever *motivated* altruistically—i.e., by a motive to benefit others rather than by selfish interests or automatic control mechanisms. Equally controversial is whether psychological altruism—the disposition to behave based on altruistic motives—could be or has been *selected for*. We propose a new evolutionary framework for investigating psychological altruism, and, on that basis, argue that some organisms have evolved by natural selection to be altruistically motivated.

In section 1, we make the questions we seek to answer more precise. In section 2, we articulate different kinds of motivation for altruistic behavior. In section 3, we outline circumstances under which different kinds of motivation for altruistic behavior should be expected to evolve. We bring out the importance of our findings in section 4. We conclude in section 5.

1. Determining When to Help

For obvious reasons, in general it is not adaptive for a given organism to either always help other organisms—or even all organisms from the same species—or never help anyone else. Organisms generally have to *determine* when to help others. Two points are important about this.

First, determining when to help others does not require that an organism deploy a concept of adaptiveness or fitness. All that is required is that the organism rely on a psychological mechanism for selecting altruistic behaviors that is based on psychological variables that correlate with the (inclusive) fitness of these behaviors. It is plausible that organisms have such a mechanism, as organisms that systematically picked out maladaptive behaviors would eventually go extinct. Furthermore, the assumption that many organisms select altruistic behaviors in ways that are by and large adaptive for them has empirical support (Houston & McNamara, 1999; Jensen, 2012; Chudek et al. 2013, 436-437).¹

Second, there are two types of mechanism by which an organism can determine when to help. On one hand, the organism could rely on *automatisms* such as reflexes or fixed action patterns—i.e., its helping behaviors might simply be triggered in appropriate situations. On the other hand, the organism could rely on *desires*—i.e., representational conative states—to produce the behavior.² For present purposes, we restrict ourselves to the latter: the discussion to follow concerns organisms that are cognitively sophisticated—i.e., organisms that rely on representational mental states when interacting with their environment (for more on this, see Schulz, 2016, 2018). This is not to say that representational decision-making is the biologically most widely instantiated (or otherwise most important) way of interacting with the environment. The point is just to focus on those forms of helping behavior for which the question of altruistic motivation arises: there is just no reason to classify automatism-driven behaviors as either psychologically altruistic or psychologically egoistic. That said, we return to some aspects of automatism-driven helping behavior below.

With this in mind, we seek to provide answers to the following two questions. Consider the different ways in which a representationally driven organism can determine whether to help another. Which ones deserve to be called *altruistic*, which *egoistic*, and which deserve neither label? We will answer this question in the next section. The subsequent section will examine which of the different mechanisms are likely to be selected for, and under which circumstances.

2. Psychological Altruism: Classical vs. Nonclassical

A common way to distinguish altruistic from egoistic motivations to help others is based on the content of the organism's *ultimate* desires—i.e., those desires that are not derived from other desires (see Sober and Wilson, 1998; Stich et al., 2010; Garson, 2016; Schulz, 2016). Organisms motivated by ultimate other-involving desires (i.e., desires directed at increasing others' wellbeing) are said to be psychological altruists. Organisms motivated by ultimate self-involving desires (i.e., desires directed at increasing one's

¹ In cases where the relevant organisms are cultural learners, what determines whether a behavior will spread through the population can depend on more than its biological adaptiveness (Boyd & Richerson, 2005; Stich, 2016).

² Desires, in the sense relevant here, need not be propositional representations of a state of affairs and need not be explicitly deliberated with: an immediate urge *to help another organism*—i.e., a representation of the kind “must help so and so”—without representing the precise state of affairs aimed at and without deliberating, is sufficient (cf. Clavier, 2011). We will remain neutral on the vexed issues of what counts as a representation and how it gets its content. Any reasonable account will do; for an opinionated defense of representational explanation within cognitive neuroscience, see Boone and Piccinini (2016).

own wellbeing) are said to be psychological egoists. (Organisms motivated by ultimate neutral desires—i.e., desires directed neither at the self nor at another organism—are thus neither altruists nor egoists.) As we show in this section, however, this characterization is too simplistic: we must also consider how desires are *produced*. A desire can be produced in ways that deserve to be called either “altruistic” or “egoistic.” To make progress on the empirical and philosophical investigation of psychological altruism, therefore, the ways in which the relevant desires are produced must be considered.

A more adequate way of distinguishing between psychological altruism and egoism considers *two* dimensions along which motivations to help others can differ. First, as the recent discussion of this topic in the literature makes clear (Sober & Wilson, 1998; Stich et al., 2010; Batson, 1991), helping behaviors might result from desires having different kinds of *contents*: organisms might be driven by a genuine concern for other organisms, or they might be driven by a concern for themselves only (or even by a concern that is neither other-involving nor self-involving, such as for equal distributions of resources). Second, helping behaviors might result from desires that are *produced* in different ways: whatever their content is, they could be produced altruistically, egoistically, or neutrally. This aspect thus concerns differences in the psychological mechanisms that generate and maintain an organism’s desires. To make this clearer, consider these differences in desire production in more detail.

A desire is *egoistically produced* if and only if it is generated by psychological mechanisms that were selected for increasing their bearer’s own reproductive success (i.e., its direct fitness) only, as opposed to that of other organisms (see also West et al., 2007).³ Two mechanisms of this kind can be distinguished. First, organisms may have an innate disposition to form desires furthering their own self-interest, which is triggered in a specific set of circumstances. That is to say, the organism’s brain may be structured so that detecting the occurrence of certain states of affairs (e.g., low blood sugar levels) engenders the formation of a desire that leads to an increase in their own wellbeing (e.g., an urge to eat). Second, organisms can learn to produce desires by conditioning: if a desire is produced and maintained because its previous instances were found to be rewarding (or if the lack of such a desire led to punishment), it is egoistically produced. For example, if past tokens of the desire in question (or reasonably similar ones) improved the agent’s emotional state (or their absence worsened it) to such an extent that this past emotional reward (or punishment) alone is sufficient to cause the desire in question in the appropriate circumstances, the desire is egoistically produced (see also Garson, 2016; Stich, 2007; and Goldman, 1970; Rachlin, 2002).⁴ In this last case, the psychological mechanism that produces the desire after conditioning is the organism’s tendency to pursue rewards and avoid punishments; for present purposes, we conservatively assume that this mechanism has been selected for only because it contributed to the organism’s own (direct) fitness. If there are cases in which the organism’s disposition

³ Put differently, egoistically produced desires are desires produced by mechanisms whose evolution is driven just by the fact that they increase the direct reproductive fitness (i.e., the number of offspring) of the organism in question, while not increasing its indirect fitness (so that the second component of the inclusive fitness calculation is zero).

⁴ Note, though, that in cases where conditioning operates by rewarding a previously occurring desire, the chain of rewards must end in a desire that is produced by some other means.

to pursue rewards and avoid punishments has also been selected because it contributed to other organisms' fitness, then those are cases of altruistically produced desires.

By contrast, a desire is *altruistically produced* if and only if it is generated by psychological mechanisms that were selected for because they increase the reproductive success of other organisms (i.e., their direct fitness) as well as, possibly but not necessarily, that of their bearer.⁵ The main such mechanism consists of altruistically evolved innate dispositions to form desires to help others, which are triggered in a specific set of circumstances. That is to say, the organism's brain may have evolved to be structured so that detecting the occurrence of certain states of affairs (e.g., a crying baby) engenders the formation of a desire that leads to an increase in the other's wellbeing (e.g., an urge to soothe the baby). (It is worth noting that the generation of this desire may need to be mediated by an intermediate internal state, such as empathy,⁶ as well as appropriate background conditions, such as bonding between agent and target and sufficient resources. This does not affect the substance of the issue here.)

Finally, a desire could be *neutrally produced*. This happens if and only if it is neither egoistically nor altruistically produced. A desire may be produced by mechanisms that have not been selected for at all, or it may be a by-product of other mechanisms. For example, desires might sometimes be produced similarly to how some skills are acquired. So, after repeatedly helping B, organism A may come to form a desire to help B, much like the skill of riding a bike comes after an organism has practiced this for a while. It is empirically unclear to what extent this possibility is actually instantiated (Schroeder, 2004; Garson, 2016); here we leave it open as a possibility.

Taking a step back, we can now put together these two aspects—the content-focused and the production-focused—of how helping behavior can be motivated. This will lead to at least three different motivational architectures underlying helping behavior:⁷

Psychological egoism selects actions based on desires with egoistic contents.⁸

⁵ We may distinguish further between two kinds of altruistically produced desires. *Strictly* altruistically produced desires are produced by mechanisms selected for increasing solely others' direct fitness—they increase the expected reproductive success of that other organism and do not affect or even decrease the bearer's expected reproductive success. *Broadly* altruistically produced desires are produced by mechanisms selected for increasing others' direct fitness along with their bearer's (what West et al., 2007, call "mutual benefit"). This distinction will not play a role in this paper.

⁶ De Waal (2008) argues that empathy can play such a role; Klimecki et al. (2016) provide additional supporting evidence. Someone might object that empathy-driven altruistic behavior is *selfishly* motivated, because it improves the agent's emotional state. This objection is confused. It is well established that empathy can lead to either altruistic or selfish desires and, consequently, to either altruistic or selfish behaviors (Schulz, 2017). Here we are considering cases in which empathy leads to ultimate altruistic desires. In such cases, empathy deserves to be considered a component in an *altruistic* source of desires. Any improvement in the agent's emotional state, which may or may not follow the desire's satisfaction, is not the agent's motive—it's just a by-product.

⁷ Böckler et al. (2016), Clavien and Chapuisat (2013), Ramsey (2016), and Garson (2016) also distinguish among different types of altruisms, but their theoretical frameworks are very different from ours.

⁸ While psychological egoism often results in egoistic behavior, it can also result in altruistic behavior.

Classical psychological altruism selects actions based on non-egoistically produced ultimate desires with altruistic contents.⁹

Nonclassical psychological altruism selects actions based on egoistically produced ultimate desires with altruistic contents.

This division among types of helping motivations makes clear that the traditional definitions of psychological altruism and egoism are oversimplified: they make it appear that there are only two motivations for altruistic behavior, when in fact there are at least *three*.¹⁰ This matters, as it can lead researchers to overlook important options that need to be considered in our theorizing about the psychological structures that might underlie altruistic behaviors. Making this clearer is the aim of the next section.

3. Psychological Altruism by Natural Selection

Are these different psychological motivations to help others merely theoretical possibilities, or is there reason to think that they have in fact evolved to drive the helping behaviors of different organisms? To answer this question, we now lay out some of the major adaptive pressures on these different ways of making helping decisions. This sort of evolutionary psychological project is associated with many uncertainties (Richardson, 2007; Buller, 2005; Garson, 2014), so its conclusions should not be overstated. Nevertheless, evolutionary analysis is a useful launch pad for investigating how helping behavior is motivated (Schulz, 2018).

1. *Psychological egoism*

Psychological egoism is the most flexible but also the most cognitively demanding way to generate altruistic desires. In some cases, it may well be the best strategy—such as when helping someone makes it very likely that the organism will reciprocate that help in the future. However, in many cases it is simply *too* cognitively demanding, and hence unfeasible. There are a large number of variables that influence whether helping another organism is adaptive (e.g., Frank, 1998; West et al., 2011; Queller, 1992; Okasha, 2006; Birch & Okasha, 2014; Skyrms, 1996, 2004; Sober & Wilson, 1998; Stevens and Hauser 2004). Consider meeting a hungry stranger; whether feeding them is adaptive depends, inter alia, on how closely they are biologically related, whether they are likely to reciprocate, whether there are any social mechanisms for rewarding sharers or punishing non-sharers, and whether such mechanisms are likely to be triggered in that circumstance. Given how many unknowns there are, calculating whether to feed a hungry stranger from egoistic first principles is generally unfeasible. In

⁹ Classical psychological altruism can, in turn, be divided into two kinds: a *pure* kind, where the ultimate desires with altruistic content are all produced altruistically (whether strictly or broadly), and an *impure* kind, where the ultimate desires with altruistic content are produced neutrally. We will not consider this subdivision further here.

¹⁰ Note also that one type of helping motivation that is not on the list is *impersonal agency*, which selects actions based on desires with neutral (i.e., non-egoistic and non-altruistic) contents. Since this kind of impersonal helping is most relevant to highly social and cognitively sophisticated creatures subject to cultural pressures, such as human beings, and presumably it piggybacks on other forms of altruism, we set it aside here.

many practical circumstances, organisms simply lack sufficient information to conduct the relevant instrumental reasoning with any hope of reaching reliable conclusions. For this reason, psychological egoism is unlikely to be the most important way of making helping decisions: it may be plausible in a restricted set of cases, but it is unlikely to be the primary explanation of helping behavior.

2. *Classical psychological altruism*

Classical psychological altruism should be expected to be an important motivator of helping behavior. The easiest way to see this is by noting that it is widely accepted that evolution can select for (innate) *automatisms* that produce altruistic behavior. Given this, there is no reason to rule out that it can also select for (innate) ultimate other-involving *desires*.

One way to see that classical psychological altruism is likely to evolve is by noting that, by definition, an organism's inclusive fitness is a positive function of the (direct) fitness of organisms that are sufficiently closely (genetically) related to the focal organism (Gardner et al., 2011; Taylor & Frank, 1996; Frank, 1998; Queller, 1992; van Veelen, 2009; Birch & Okasha, 2014). Therefore, other things being equal, it is adaptive for many organisms to help sufficiently close kin when they are in need. This is not true for all organisms—it depends on the exact nature of the benefits and costs involved, but it is true for a number of organisms, including many mammals (Gardner et al., 2011; Taylor & Frank, 1996; Frank, 1998).¹¹ Of course, for some organisms, helping their kin can be done automatically, without involving desires at all (see e.g., Strassman et al., 2011; Kuzdzal-Fick et al., 2011). In cognitively sophisticated, representation-driven organisms, however, this circumstance creates adaptive pressures to have an innate disposition to form a *desire* to help kin when they are in need. As noted above, this is a form of classical psychological altruism: it is an altruistically produced desire with the content to help another organism. Moreover, there are probably other examples of classical psychological altruism as well: in some populations of organisms, similar reasoning might well support helping injured in-group members, at least in some cases (Sober & Wilson, 1998). For present purposes, the key point to note is just that there are important sets of circumstances in which classical psychological altruism can be expected to be adaptive.

3. *Nonclassical psychological altruism*

Nonclassical psychological altruism is also likely to play a large role in animal psychology. Like classical psychological altruism, nonclassical psychological altruism is less cognitively demanding than egoism: it need not derive all helping behaviors from egoistic first principles. Unlike classical psychological altruism, however, nonclassical psychological altruism does not depend on the fact that the desire to help certain other organisms is cross-generationally adaptive. Instead, it can allow for the adaptiveness of this desire to be dependent on environmental factors, such as the likelihood of reciprocation. Put

¹¹ This may require complex decisions as to which kin to help (also taking into account potential future kin). See also Hausfater & Hrdy (1984) and Trivers (1974).

more positively: nonclassical altruism is adaptively favored whenever *learning* whether to want to help others is adaptive, even though calculating whether to do so from egoistic first principles is unfeasible.

For example, nonclassical altruism is adaptive if it is inter-generationally variable whether helping is adaptive, but intra-generationally stable. For example, many social animals form alliances with select members of their group to share resources and maintain social status. In cases like this, it is not adaptive for (representationally-driven) organisms to be born with an innate disposition to want to help certain other organisms: after all, whether helping is adaptive depends on the precise conditions the organism faces. However, if the conditions are such that helping *is* adaptive, it is also adaptive for the organism not to derive the helping behavior, every time, from egoistic ultimate desires. The existence of circumstances that fit this general pattern is well known: much of gene-culture coevolutionary theory is dedicated to investigating these circumstances (see e.g. Henrich, 2015; Boyd & Richerson, 2005; Fehr & Fischbacher, 2003). In a nutshell: given that it is sometimes adaptive for organisms to learn when to *behave* altruistically, there will likely be circumstances where it is adaptive for organisms to learn to be *motivated* to behave altruistically—i.e., to be nonclassical altruists.

4. The Importance of Psychological Altruism for Other Disciplines

These conclusions have several important consequences for the many disciplines studying altruism. For starters, the *moral* value of psychological altruism vis-à-vis egoism must be reassessed. This is not only due to the fact that altruism and egoism can equally reliably lead to the same kinds of helping behaviors (as has been argued by Stich et al., 2010), but also because even behaviors caused by desires with other-regarding contents might be egoistically produced. This matters, as it is not obvious that there is anything morally problematic about nonclassical psychological altruism. Therefore, to see “the basic goal of morality as ‘selflessness’” is too strong (see Schroeder, 2000, p. 396; Rachels, 2000, p. 81): moral action can countenance some egoistic influences on helping behavior.

Further, a better understanding of the differences between egoism, classical altruism, and nonclassical altruism matters, as it may help us determine how to create a more cooperative global culture. For example, knowing more about how we can learn to be altruists (e.g. through training our empathy mechanisms to recognize *all* human beings as in-group members) might help lessen the prevalence and severity of the kinds of racist and discriminatory behaviors that are still so common (Greene, 2013; Klimecki, 2015).

In cognitive neuroscience, recent work has emphasized that understanding helping behavior requires investigating its underlying neurocognitive structures (Gluth and Fontanesi, 2016, Greene et al., 2016, Hein et al., 2016, Kurzban et al., 2015). We wholeheartedly agree with this. Our conclusions show that doing so requires a diachronic approach: we need to take into account not only how the organism is constituted at time t_1 (by assessing the contents of its desires), but also how it was constituted at time t_0 (by assessing how it produced the relevant desires).

Finally, our argument also complements recent work in economics. It is a still typical (though not theoretically required) assumption in much economic modeling that people are egoistically motivated. This assumption has come under much empirical and theoretical criticism, especially from work in

behavioral economics and neuroeconomics (Rand, 2016; Fehr & Camerer, 2007; Falk et al., 2003; Clavien and Chapuisat, 2016). Our present argument adds evolutionary biological considerations to the body of evidence speaking against egoistic models of human economic behavior.

5. Conclusion

There are two importantly different kinds of altruistic motivations: *classical psychological altruism*, which generates ultimate desires to help other organisms at least partly for those organisms' sake, and *nonclassical psychological altruism*, which generates ultimate desires to help other organisms for the sake of the organism providing the help. Classical psychological altruism is adaptive if the desire to help others is intergenerationally reliable, and thus need not be learned. Nonclassical psychological altruism is adaptive when the desire to help others is adaptively learnable. Thus, both kinds of psychological altruism are likely to be selected for. This theory of the motivational structures underlying helping behaviors opens up new avenues for the productive, interdisciplinary study of psychological altruism.

References

- Batson, D. (1991). *The Altruism Question: Toward a Social-Psychological Answer*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Birch, J., & Okasha, S. (2014). Kin Selection and Its Critics. *BioScience*.
- Böckler, A., Tusche, A., & Singer, T. (2016). The Structure of Human Prosociality: Differentiating Altruistically Motivated, Norm Motivated, Strategically Motivated, and Self-Reported Prosocial Behavior. *Social Psychological and Personality Science*, 7(6), 530-541.
- Boyd, R., & Richerson, P. (2005). *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.
- Buller, D. (2005). *Adapting Minds*. Cambridge, MA: MIT Press.
- Chudek, M.; Zhao, W., and Henrich, J. (2013). Culture-Gene Coevolution, Large Scale Cooperation, and the Shaping of Human Social Psychology. In K. Sterelny, R. Joyce, B. Calcott, and B. Fraser (eds.). *Cooperation and Its Evolution*. Cambridge, MA: MIT Press, pp. 425-457.
- Clavien, C. (2011). "Altruistic Emotional Motivation: An Argument in Favour of Psychological Altruism." in K. Plaisance & T. Reydon (eds.), *Philosophy of Behavioral Biology. Boston Studies in Philosophy of Science, Volume 282*. Springer, 275-296.
- Clavien, C. & Chapuisat, M. (2013). "Altruism across disciplines: one word, multiple meanings." *Biology & Philosophy*, 28 (1), 125–140.
- Clavien, C. & Chapuisat, M. (2016). "The evolution of utility functions and psychological altruism." *Studies in the History and Philosophy of the Biological and Biomedical Sciences*, 56, 24-31.
- de Waal, F. B. M. (2008). "Putting the Altruism Back into Altruism: The Evolution of Empathy." *Annual Review of Psychology*, 59: 279-300.
- Falk, A., Fehr, E., & Fischbacher, U. (2003). On the Nature of Fair Behavior. *Economic Inquiry*, 41(1), 20-26.
- Fehr, E., & Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn Sci*, 11(10), 419-427.
- Fehr, E. & Fischbacher, U. (2003). The nature of human altruism. *Nature*, " , 785-791
- Frank, S. A. (1998). *Foundations of Social Evolution*. Princeton: Princeton University Press.
- Gardner, A., & West, S. A. (2010). Greenbeards. *Evolution*, 64(1), 25-38.
- Gardner, A., West, S. A., & Wild, G. (2011). The genetical theory of kin selection. *J Evol Biol*, 24(5), 1020-1043.
- Garson, J. (2014). *The Biological Mind: A Philosophical Introduction*. Routledge.

- Garson, J. (2016). Two Types of Psychological Hedonism. *Stud Hist Philos Biol Biomed Sci.*, 56, 7-14.
- Gluth, S., & Fontanesi, L. (2016). "Wiring the Altruistic Brain." *Science* 351(6277), 1028-1029.
- Godfrey-Smith, P. (2008). Varieties of Population Structure and the Levels of Selection. *British Journal for the Philosophy of Science*, 59, 25-50.
- Goldman, A. (1970). *A Theory of Human Action*. Princeton: Princeton University Press.
- Grafen, A. (2006). Optimization of Inclusive Fitness. *J Theor Biol*, 238, 541-563.
- Greene, J. D. (2013). *Moral Tribes*. Penguin.
- Greene, J. D., Morrison, I., & Seligman, M. E. P. (2016). *Positive Neuroscience*. Oxford: Oxford University Press.
- Hausfater, G., & Hrdy, S. B. (Eds.). (1984). *Infanticide: Comparative and Evolutionary Perspectives*. Chicago: Aldine Transactions.
- Hein, G., Morishima, Y, Leiberg, S., Sul, S., & Fehr, E. (2016). "The Brain's Functional Network Architecture Reveals Human Motives." *Science* 351(6277), 1074-1078.
- Henrich, J. (2015). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton, NJ: Princeton University Press.
- Houston, A. I., & McNamara, J. M. (1999). *Models of Adaptive Behaviour: An Approach Based on State*. Cambridge: Cambridge University Press.
- Jensen, K. (2012). Who Cares? Other-Regarding Concerns—Decisions with Feeling. In P. Hammerstein & J. R. Stevens (Eds.), *Evolution and the Mechanisms of Decision Making* (pp. 299-317). Cambridge, MA: MIT Press.
- Klimecki, O. M. (2015). "The Plasticity of Social Emotions." *Social Neuroscience*, 10(5): 466-473.
- Klimecki, O. M., Mayer, S. V., Jusyte, A., Scheeff, J. & Schönenberg, M. (2016). "Empathy promotes altruistic behavior in economic interactions." *Scientific Reports*, 6: 31961. DOI: 10.1038/srep31961.
- Kurzban, R., Burton-Chellew, M. N., & West, S. A. (2015). "The Evolution of Altruism in Humans." *Annual Review of Psychology*, 66: 575-599.
- Kuzdzal-Fick, J. A., Fox, S. A., Strassmann, J. E., & Queller, D. C. (2011). High relatedness is necessary and sufficient to maintain multicellularity in *Dictyostelium*. *Science*, 334, 1548-1551.
- Okasha, Samir (2006). *Evolution and the Levels of Selection*. Oxford: Oxford University Press.
- Okasha, Samir, "Biological Altruism", *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2013/entries/altruism-biological/>>.

- Queller, D. C. (1992). Quantitative Genetics, Inclusive Fitness and Group Selection. *American Naturalist*, 139, 540-558.
- Rachels, J. 2000. Naturalism. In H. LaFollette (ed.), *The Blackwell Guide to Ethical Theory*. Oxford: Blackwell Publishing, pp. 74-91.
- Ramsey, G. (2016). Can altruism be unified? *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 56, 32-38.
- Rand, D. (2016). Cooperation, Fast and Slow: Meta-Analytic Evidence for a Theory of Social Heuristics and Self-Interested Deliberation. *Psychol Sci*.
- Rachlin, H. (2002). "Altruism and Selfishness." *Behavioral and Brain Sciences*, 25, 239–296.
- Richardson, R. (2007). *Evolutionary Psychology as Maladapted Psychology*. Cambridge, MA: MIT Press.
- Schroeder, T. (2004). *Three Faces of Desire*. Oxford: Oxford University Press.
- Schulz, A. (2016). Altruism, Egoism, or Neither: A Cognitive-Efficiency-Based Evolutionary Biological Perspective on Helping Behavior. *Stud Hist Philos Biol Biomed Sci*, 56, 15-23.
- Schulz, A. (2017). The Evolution of Empathy. In H. Maibom (ed.). *Routledge Handbook of the Philosophy of Empathy*. London: Routledge.
- Schulz, A. (2018). *Efficient Cognition: The Evolution of Representational Decision Making*. Cambridge, MA: MIT Press.
- Skyrms, B. (1996). *Evolution and the Social Contract*. Cambridge: Cambridge University Press.
- Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure*. Cambridge: Cambridge University Press.
- Sober, E. (2001). The Two Faces of Fitness. In R. Singh, D. Paul, C. Krimbas & J. Beatty (Eds.), *Thinking about Evolution: Historical, Philosophical, and Political Perspectives* (pp. 309-321). Cambridge: Cambridge University Press.
- Sober, E., and Wilson, D. S. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
- Stevens, J. R. and Hauser, M. D. (2004). "Why be nice? Psychological constraints on the evolution of cooperation." *Trends in Cognitive Sciences*, 8 (2) 60–65.
- Stich, S. (2007). Evolution, Altruism and Cognitive Architecture: A Critique of Sober and Wilson's Argument for Psychological Altruism. *Biology and Philosophy*, 22, 267-281.
- Stich, S. (2016). "Why there might not be an evolutionary explanation for psychological altruism." *Studies in the History and Philosophy of the Biological and Biomedical Sciences*, 56, 3–6.

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

Taylor, P. D., & Frank, S. A. (1996). How to Make a Kin Selection Model. *J. Theor. Biol.*, 180, 27-37.

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

van Veelen, M. (2009). Group Selection, Kin Selection, Altruism, and Cooperation: When Inclusive Fitness is Right and When It Can Be Wrong. *J Theor Biol*, 259, 589-600.

West, S. A., Griffin, A. S., & Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J Evol Biol*, 20(2), 415-432.

West, S. A., El Mouden, C., & Gardner, A. (2011). Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Human Behavior*, 32(4), 231-262.