

The Evolution of Psychological Altruism

Student Name: Nico Heidari Tari

Student Number: s1026675

Supervisor: Bart Geurts

Radboud School of Philosophy

Radboud University Nijmegen

Thesis for obtaining a “Master of arts” degree in philosophy.

15 June, 2020

I hereby declare and assure that I, Nico Heidari Tari, have drafted this thesis independently, that no other sources and/or means other than those mentioned have been used and that the passages of which the text content or meaning originates in other works - including electronic media - have been identified and the sources clearly stated. Place: Arnhem. Date: 08-06-2020

Introduction

Imagine a man sitting at a train station eating his lunch. Suddenly, a pigeon walks by and the man decides to share some small chunks of bread with the animal. This means quite a lot for the bird, since it can likely subsist on that food for many hours. The donator, on the other hand, only misses part of his lunch and does not really get much out of the transaction, perhaps only the pleasant knowledge that he has helped another sentient being. Humans do this sort of thing all the time. We share food and knowledge, help each other move, donate clothes to the poor, and so on without directly expecting something in return. *Prima facie*, it appears that many of our acts are altruistic.

A substantial amount of research has been undertaken to investigate exactly why such behaviors exist, and why they are relatively common, since from an evolutionary perspective, altruism is not a successful strategy for survival. One can find altruism in the animal kingdom in specific types of situations, and human altruism in particular appears to be something *sui generis* and quite prevalent. It might seem that an explanation for this is that our conscious human minds are capable to exercise agency and travel beyond the boundaries put on us by biology. As such, they would not be able to be explained with an evolutionary theory. It is rather my viewpoint that there is an evolutionary explanation behind human altruism, and in this essay I aim to provide this.

Thus, the main question I aim to answer in this paper is: does altruism in humans have an evolutionary explanation? Later in this paper I answer this question affirmatively. Humans are unique beings on planet earth, since our culture allows us to interact with others and the environment in incredibly complex ways. It is possible to assume that since culture is such an enormous force in our lives, that somehow makes us rise above our biology, and allows us to do things, like giving food away for free, that would not be appreciated by purely Darwinian principles. Even though this explanation is viable to explain altruistic acts in humans, I think it is not the complete picture. What it misses is the biological dimension in culture: culture is influenced (significantly) by our biology, and in turn it exercises power over biology and changes it as well. The main point behind the evolutionary explanation of human altruism is that culture interacts with biology, and is thus subjected to certain evolutionary forces.

I answer my research question in the following manner. I first start out by providing a theoretical overview of the concept of altruism. This concept has had a long and troublesome history, and I discuss the most important insights researchers have gained that are relevant to my essay. I use the theory to establish a distinction between two different forms of altruism, namely: biological and psychological altruism. Then, I argue that the distinction is not so

clear cut after all, since psychological altruism still depends on our biology too. A concept in my argument is the so-called Baldwin effect. This effect ensures that culturally learned behaviors can become incorporated in our genetic makeup. In the next section, I elaborate on the interaction between culture and nature that produced humans that are designed behave altruistically. This interactive theory is called culture-gene coevolution, and I show that this theory is also applicable to explain altruism in humans. I continue with a discussion of Stich's (2016) critique of evolutionary explanations of psychological altruism, providing counterarguments in support of the view that psychological altruism has influences from evolution. Finally, I end with a brief conclusion with the main points of this paper.

Theoretical Background of Altruism

The notion of altruism is puzzling from an evolutionary perspective. In Darwin's (1859) original writings on natural selection, organisms are inclined to maximize their reproductive success or 'fitness'. If they wouldn't, their genes would be wiped out in due time. So, if humans help each other all the time, and sometimes don't expect anything in return, then this does not appear to be an evolutionary viable strategy. This is because the altruists in a population are susceptible to 'cheaters' who can reap the social benefits while giving nothing or barely anything in return (Boyd & Richerson, 2006; Travisano & Velice, 2004). Such cheaters will then thrive and maximize their fitness, thus overshadowing the altruists across generations. The fate of the altruist looks dim, as he or she will likely perish in the long run.

So, how is it possible that there are so many individuals walking around, who are driven to do good things without expecting much in return? A number of theories have been put forward to explain such cases. One important insight was Hamilton's rule (Hamilton, 1964). According to this rule, individuals will cooperate and help other beings when the cost of the transaction is less than the benefit, weighted by the relatedness between the individuals. So, for example, we will go to great pains to help out our children, who share around 50% of their genes with their parents. But doing the same favors for a second cousin seems more inappropriate, perhaps only occurring in the rarest of circumstances. Hamilton formulated this rule in an elegant formula which states that a certain social behavior will be selected for if and only if $rb > c$, where r stands for the relatedness coefficient between the individuals involved¹, b is the generated benefit for the beneficiary, and c is the cost for the benefactor. This theory is also known as 'kin selection'.

¹ The relatedness coefficient r is traditionally interpreted as a probability. However, it is sometimes also reformulated as a regression: as a relationship between variables. Regressions can also be negative, unlike

In kin selection, researchers focus on a concept known as *inclusive fitness*. With regular biological fitness, one would look exclusively at the number of offspring a single individual could produce. With inclusive fitness however, the number of offspring that individual can help survive with various helping behaviors is added to it. So, if a certain animal shows altruistic behaviors, it might be an unsuccessful strategy if we look at individual fitness, but it might be very successful if we look at inclusive fitness: it helps a significant portion of their genes to survive.

This was a huge step in the explanation of altruistic acts that seemingly do not make sense from an evolutionary perspective. However, it still leaves many cases of human altruism unsolved, since humans do many helpful things for unrelated people too, or even animals. The example of the man in the train station is a case in point. Some theories have been put forward to explain the remaining gap in human cooperative behavior.

One such theory is called reciprocal altruism by Trivers (1971), which states that animals enter in *quid pro quo* relationships with each other, where they help each other out and expect similar favors in return. For example, chimpanzees regularly groom each other (Gomes, Mundry, & Boesch, 2009; Newton-Fisher & Lee, 2011) and are often reciprocated, even over long stretches of time. In reciprocal altruism, helping behaviors can be selected for if the cost of the altruistic act is outweighed by the likelihood of getting repaid. This makes it possible for individuals who are not related, but do interact frequently, to establish relationships where they can act altruistically.

Another theory that has gained traction is group selection (Chudek et al., 2011; Sober & Wilson, 1999), where the functional unit of evolution is taken to be the group, rather than the individual. This was a rather controversial topic, since the scientific consensus was that evolution acted on the individual (Dawkins, 1976; Hamilton, 1964; Smith, 1964). Even though it had some stigma to overcome, group selection theory has been revised and has returned as a viable explanation for certain phenomena. Nowak, Tarnita and Wilson (2010) argued that group selection can occur when competition between groups puts higher pressure on survival than competition between individuals. They view selection pressure as a layered process, where it can act on the group, individual, and even on the underlying cells. This theory is referred to as multi-level selection.

probabilities, and so negative behaviors such as harm can be explained with this model. Both interpretations are fine for this paper, because the focus is on altruistic behaviors, and so harmful behaviors are not taken into consideration.

An interesting result found by Marshall (2011) is that inclusive fitness and group selection are formally equivalent (p. 329). Thus, in mathematical terms, one can be reduced into the other. He states that these two theories might just be two different perspectives on the same process: traits being passed down from generation to generation, and the most successful ones surviving. One theory just looks at the individual, and the other one at groups. Thus, the reader should keep in mind that whenever I discuss one of these theories, even though the concepts appear different, they might have an equivalent structure underneath.

Biological Altruism and Psychological Altruism

The aforementioned theories indicate that research on altruism has been quite fruitful. Many cases of altruism can be explained so far. For example, some social insects like bees have sterile workers who cannot reproduce. In terms of fitness, this is the greatest altruistic act an individual could do, because her individual genes will certainly not survive, while she is helping others to pass on theirs. This phenomenon can be explained with kin selection theory, because the relatedness between workers is abnormally high due to their sex-determination system called haplodiploidy. Since their relatedness coefficient r is high, this can call for extreme acts of altruism, if the benefits are high enough.

Also, many cases of altruism that are between individuals who are not related make more sense now. For instance, vampire bats often regurgitate blood to conspecifics, so that they can prevent them from dying (Wilkinson, 1984). These type of bats only consume blood to survive, and they can go on for approximately 70 hours without a meal. If one vampire bat had a successful hunt but finds another member of his group who was unlucky for two whole days, he can share some food to help out. This might mean that he can survive a day shorter, but at least the other conspecific does not die. Thus, there is a low cost and a high reward, in such cases they can go for it. And the data (p.183) shows that altruistic individuals who helped out in the past are more likely to be reciprocated, thus being a convincing case for the theory of reciprocal altruism.

These theories explain many cases we can find in nature. They are also applicable to humans to a certain extent, since we are also biological organisms that need to survive in our environment, and make decisions about how helpful we want to be to the individuals around us. Our energy is limited, so we cannot help everyone equally. Kin selection explains why we do so much for our children and close family members. Reciprocal altruism clarifies why we like to have close friends, where we continuously do favors for each other. But there are still some cases that remain unresolved. Why would we give food to a random animal on the

street, who is not even in the same species as us? Especially when we know that reciprocation is not an option.

Even though theories like kin selection and reciprocal altruism can explain some forms of human altruism, philosophers have argued that in human cases, there is one added dimension that makes our acts essentially different than in nonhuman animals. This extra dimension is *motivation*. If a human does a right act but is motivated in a very selfish manner, we tend to judge the act more negatively. For example, if a person gives money to a charity, only because he wanted to get publicity for him or herself, then we would draw the conclusion that he was not acting altruistically, but in fact selfishly. Things would be even worse if he boasted about it on social media. This is in contrast to altruism in nonhuman animals, where motivation seems to be neglected altogether, and where the focus is more on outcomes. For instance, when trying to explain why bees sacrifice their own life to save the hive from potential intruders, researchers provide explanations in terms of population structure and fitness (Blows & Schwarz, 1991; Mullen & Thompson, 2015). So, these explanations are in accordance with the theories of kin selection and reciprocal altruism.

The more human type of altruism which focusses more on motivations or intentions, rather than consequences is called *psychological altruism* by philosophers, biologists and psychologists (Ananth, 2005; De Waal, 2008; Ramsey, 2016). More specifically, altruistic acts need to be motivated by so-called ‘ultimate desires’ for the well-being of other people (Doris et al., 2020). An ultimate desire is a sort of bedrock desire in the context of a specific goal. For example, one of my goals might be to lead a healthy life. This is something that I desire and it might manifest itself in particular ways in day to day life. On a certain day, I might develop the desire to go to the gym because it was too long ago, or I might desire to eat a salad. But ‘eating a salad’ and ‘going to the gym’ are not ultimate desires in themselves, because they only serve instrumentally to fulfill my bigger purpose, namely leading a healthy life. Accordingly, those smaller desires have been named *instrumental desires*. Philosophers have posited these ultimate desires to avoid an infinite regress: not all desires can be instrumental, otherwise desiring in itself becomes rather meaningless (Doris et al., 2020). Therefore, as the argument goes, some desires must exist that are ultimate, that can be wanted purely for their own sake.

So, there appears to be a discrepancy in the meaning of concepts we use to explain certain behavior in the world. When looking at the human, social world, ‘altruism’ looks at motivations and desires. When looking at the biological, natural world, the same concept purely looks at fitness. This is rather peculiar, because there is no *a priori* clear reason why it

should be different. If we look at cells in an organism, the concept of a cell does not change when we look at a human or a chicken. A cell remains a small building block of life, consisting of a cytoplasm within a membrane, incorporating many biomolecules like proteins. While the concept of altruism focusses on entirely different things depending on the organism. Why is this so? How can this discrepancy be explained?

The Baldwin Effect

The distinction in this debate between the biological and the psychological world reminds one of a lingering Cartesian dualism. The mental states of humans are perceived as something special and somehow transcending of its underlying biology. It is rather my view, that the psychological is still rooted in the physical world, though it has substantially drifted away from it. In what follows I will argue why that is the case.

One important point put forward by Sober (1993) is that human minds have the active capacity to go against the demands put on us by evolution. For example, we are wired to desire fatty foods with a lot of sugar or salt. This is because such foods were scarce and highly valuable in the past. Nowadays we can find it in abundance, and we can use reason to limit ourselves and resist temptations. Thus, according to Sober, it seems likely that psychological altruism is somehow separate from the biological domain, as it has this element of an active mind shaping decisions. Even though I agree with Sober that the human mind adds an extra dimension to altruism, I disagree that this should entail that psychological altruism is somehow essentially different from biological altruism. Our thoughts, desires and beliefs are still strongly influenced by evolution.

For this, consider the Baldwin effect. This is a theory first proposed by Baldwin (1896) and then taken up by many other scholars researching evolution (Ananth, 2005; Turney et al. 1996; Weber & Depew, 2003). It is about the effects learned behavior can have on genetic evolution. This occurs when a certain type of behavior proves to be so useful for survival, that it becomes naturally selected for. When an environment changes, individuals in that environment need to adapt flexibly to the changes. This requires them to behave in a certain way, that allows them to survive further. Individuals who are genetically predisposed to exhibit the right kind of behaviors will be more successful than the individuals who don't. As time passes on, more and more individuals with the successful behavior will be in the population, until almost everyone does it. These behaviors will be so easy to acquire that it will appear as an instinct to the individual (Dennett, 2017).

Note, that the Baldwin effect is starkly different from the more controversial *Lamarckian evolution* which states that offspring directly inherit acquired characteristics from their parents (Koonin & Wolf, 2009). With the Baldwin effect, there is no such direct inheritance, only the increase of dispositions to learn certain behaviors or characteristics. Organisms are naturally selected to have the right genetic makeup, that allows them to acquire new learned behaviors to survive in a changing environment.

An example or two might be in order to fully clarify the Baldwin effect. Firstly, consider the phenomenon of lactose (in)tolerance. Typically, mammals stop producing enzymes that break down lactose right after infancy. This means that they will have many difficulties digesting milk, which can result in stomach pains and similar unwanted pains. However, in human populations where dairy farming was common and had a long history, lactose intolerance actually went down significantly (Ananth, 2005). These are, for instance, populations like the Danes and Hungarians, where the enzymes that break down lactose are produced beyond the years of infancy. This is an instance of genes coding for changes in the body, caused by behavioral changes certain populations made.

Furthermore, as a more general example, suppose that in some animal species it happens frequently that an alpha male becomes dominant and tries to hoard all the females for himself. This entails that most of the less dominant males will have a much smaller chance to reproduce and pass on their genes. But instead of giving up, suppose two beta males team up to take on the alpha. Together, they can seriously harm the alpha or even kill it, and thereby get access to reproduction again. As this happens time and again across many generations, the tendency or disposition to display this type of cooperation goes up by means of genetic selection. At some point, cooperation to accomplish certain goals will appear to the individual as an instinct, since it is such an evolutionary successful strategy that is selected for by the genes.

This process ensures that cooperative traits will continue to exist and become more engrained in a particular organism. In the hominid lineage, cooperation was already in place in various forms, for example with food sharing, coordinated hunting, or reciprocal grooming. At some point, language started to emerge, and we were able to verbalize the importance of such behaviors. We were already aware of it in a bodily, evolutionary sense (since we were already doing it), but then we also learned how to express it into words. As such, it became part of our culture. Nowadays, our societies have become very complex, but we still hear the echoing of how important it is to work together. Collaborative efforts are encouraged and praised in science, politics, business and so on. These cultural practices are influenced by our

biology: they had more basic forms in (pre)history, and got the opportunity to get amplified through natural selection.

So, returning back to the example of the man in the train station, it seems tempting to say that his individual mind was the sole agent that decided to share food with another animal. This mind is then somehow above the Darwinian laws of biology, and can decide to break the law every now and then. Again, this perspective on the situation is permeated by Cartesian dualism.

A much more plausible explanation is that such altruistic behaviors have evolved through processes of natural selection. The Baldwin effect ensures that prosocial helping behaviors that are deemed useful by a group are praised and reinforced. Thus, a selection pressure arises where individuals who are more altruistic fare better in their environment. They would do better because they receive positively reinforcing feedback from their group members, such as status and respect. They would also be able to better cooperate with their group members due to their helpful nature.

Good cooperation in early hominids was a sure-fire recipe for survival. The environment was changing drastically around 6 million years ago, and this meant that cooperation would become the key to survival. After several generations have passed on, and natural selection has done its work, certain altruistic behaviors like food sharing will become so natural and easy to learn that it feels like an instinct for a human. It has even evolved so far right now that things like food sharing have become a pleasurable activity that is something good in its own right. Thus, human psychological altruism has an evolutionary explanation after all. This is a position that I share with Ananth (2005), who also thinks that psychological altruism is a result of successful behaviors being selected for through the Baldwin effect.

Even though our positions are rather similar in this respect, I would take this even one step further and argue that not only culture has an influence on which genes will succeed in the environment, but also the other way around. Our genes are responsible for how our bodies are built. Thus, the fact that we are built with a certain cognitive structure allows us to store and transmit cultural knowledge at a vastly accelerated rate (Sasaki, 2013). With this, one can also think of the fact that humans are so apt at acquiring language. And language perhaps can be quite accurately regarded as a system to store and transmit knowledge. So, culture and nature actually are in a bidirectional relationship with one another. They are both constantly influencing each other. This will be further elaborated upon in the next section.

Norm-Psychology and Acquisition of Social Rules

The bidirectional relationship between culture and nature have been referred to as the culture-gene coevolutionary theory (Chudek et al., 2013). This theory states that there exist two evolutionary processes, namely cultural evolution and genetic evolution. These are two separate domains, they both evolve in their own way. But they interact with each other continuously, forming a feedback loop (Laland, 2008). Culture has its influence on genes and vice versa. Our genes build us with psychological endowments that dispose us to acquiring cultural rules, behavior and motivations.

One aspect of these psychological predispositions is that we own a certain norm-psychology: a special type of psychological makeup that allows us to learn certain rules faster. Chudek et al. (2013) state that people have intuitive assumptions about social rules: that they exist, ought to have an impact on behavior, and that failure to follow these rules most likely will have repercussions. Rakoczy et al. (2008) have provided empirical evidence that 2- and 3-year old children already understand norms in contextual settings, and that they reinforce these norms by sanctioning violators. Research on young children is a great way to find clues on how humans are structured instinctively, because they haven't had many strong influences by culture yet. Their language is also rather primitive, so much explicit teaching about complex social rules will probably not be understood. Therefore, having 2- or 3-year-olds grasp social norms with ease, seems to imply that they are already cognitively structured for the acquisition of social rules.

Furthermore, 1-year-olds already show a propensity to help those in need (Warneken, 2016). For example, if an adult experimenter drops an item which is out of reach, children will pick it up and return it without any explicit instruction. Children that young do require the scene to be somewhat obvious, for example, the experimenter actively trying to reach for an object. However, when children reach their second birthday, they can help out even without any behavioral cues. For instance, they can return a dropped item, even if the adult did not notice it fell in the first place (Warneken, 2013). These experiments show that very young children are cognitively quite advanced, since they can understand goals and desires of other agents, and use contextual information to bring about useful changes in the environment.

But what motivates these individuals? One might make the objection that children portray helping behaviors for selfish reasons, like expecting to be praised or gaining rewards after they helped out. Studies have shown that this is not the case, and that children cooperate in a way that we might call altruistically motivated. That is to say, that they help out another individual for its own sake, as an activity that deserves to be pursued without it instrumentally serving to further some other goal.

Research that has yielded helping behaviors in children (Dunfield et al., 2011; Warneken & Tomasello, 2007) did not offer rewards of any sorts to the test subjects. Thus it is not necessary at all to offer something in return for helping behaviors to emerge. In fact, one study has shown that offering rewards might even cause detriment to helping behaviors (Warneken & Tomasello, 2008). This can get in the way of the intrinsic motivation of the child to help, and thus change the relationship in something revolving around direct costs and benefits. It basically turns into a business relationship, instead of two human individuals trying to be nice to each other. In the past, it has been argued that young children will only help if they are offered some reward in return (Cialdini et al., 1981). But the findings of Warneken & Tomasello (2008) go firmly against this claim.

Moreover, does the perception of parents play a role, that children are under the impression that they have to comply by the rules, or otherwise may face disciplining actions of sorts? That is not the case, as experiments show that children engage in helping behaviors even when the parents are not present (Warneken & Tomasello, 2013). Thus, helping behaviors that are psychologically altruistic are already present in young children, between the ages of 1 and 3. They are intrinsically motivated to help for its own sake, rather than relying on external rewards. So, according to the established philosophical terminology, we can conclude that infants already have ultimate desires to help fellow individuals. Our evolved norm-psychology ensures that such behaviors attain a normative structure, which allows individuals to believe that there is just something ‘right’ about acting in such a cooperative manner, and thus that it is also legitimate to punish other who ‘cheat’ or defect in some way. Children are already endowed with such a norm-acquiring structure, and it develops further as they grow up.

Additionally, more evidence supporting the claim that we are endowed with a certain norm-psychology comes from neuroeconomics: the interdisciplinary field of research that tries to understand the interaction between human decision making and its neuronal structure. Research has shown that when people adhere to social norms, that reward centers get activated in the brain (Tabibnia et al., 2008). This confirms why it feels good to follow social rules.

An interesting question that could be asked here, is whether or not following norms and doing good things for others is, at the end of day, selfish. It could be argued that we only do good things because it makes us feel good, similar to eating a nice meal or watching our favorite show. I think the line is a bit blurry here, it is difficult to say if it is clear cut selfishness or altruism. Perhaps a blend of the two concepts would be most appropriate here

(altruism). But what is in any case more certain, is that our specific norm-psychology is naturally selected for. It is this type of psychology, predisposed to follow rules and be helpful to other beings, that was most successful in ancient times. It has now become instinctual to adhere to social norms, and because of our norm-psychology, being helpful to others became a good in its own right.

With this additional information we can perhaps better understand the warm ‘fuzzy’ feeling the man at the train station might get when sharing his food with the pigeon. As mentioned before, food sharing has been an important activity that ensured the survival of our species. As such, it became a normative practice: one should share food if the cost for the benefactor is low, and the gain for the recipient high. Early humans who failed to cooperate may have been sanctioned, or even ostracized from the group. But the humans who did were respected and gained social status. With enough time, the Baldwin effect can establish genetic affects for this type of conduct, for example by enhancing reward structures in the brain to improve the likelihood of individuals displaying such behaviors. Thus, the man at the train station is predisposed to display such helping behavior, even though the animal is not a member of his own species. Our specific norm-psychology ensures that it feels good to help another, and this has developed so far that it even allows us to help other animal species.

For human psychological altruism to emerge, it is foundational that certain social norms are in existence and enforced (Warneken, 2016). Other researchers argue that socialization in this manner can even be called necessary (Boyd & Richerson, 2006). But recall, that this process of acquiring and enforcing social norms has a biological dimension to it, namely that our genetic makeup has endowed us with an inclination towards learning and enforcing social rules.

A Critique and a Reply

So far I have argued that acts of psychological altruism are the result of culture-gene coevolutionary processes. They did not appear *de novo* as a result of a disembodied cognizing mind. Rather, our specific genetic makeup disposes us to want to do good unto others. Through the Baldwin effect, we are imbued with a norm-psychology that predisposes us to be cooperative and actively seek out to help others. This behavior was fitness enhancing for our species in the past, and it kept on developing by a culture-gene coevolution.

Now, philosopher Stich (2016) has argued that psychological altruism might not have an evolutionary explication after all. He states that while experiments have shown that psychological altruism exists in some cultural groups, this is the exception rather than the

rule, and it does not mean that such altruism is something more universally rooted in biology. More specifically, he refers to a series of experiments performed by Batson et al. (2003) and Batson (2014), where in some cases test subjects displayed altruistic acts. But Stich is still not convinced, because the participants were only a small subset of the entire human population, namely, they were: Western, educated, industrialized, rich and Democratic (which has the acronym: WEIRD). Thus, the test subjects are not representative for the wider population.

Henrich et al. (2010) were the first ones to raise the general point that much empirical researchers in the West have sampled from this specific subpopulation. It was assumed that WEIRD people were representative enough to make generalizations about the wider population. It turns out that this subpopulation is more of an outlier and shows lots of variation with other cultures. Hence, many claims in sociology and psychology about human nature that were thought to be universal have become suspect. More cautious claims should be made that only generalize to the WEIRD people, and to make statements about other cultures, those should be sampled and scrutinized carefully.

So, Stich (2016) uses this general principle to argue that it is also applicable to empirical data on psychological altruism. It has been found, but predominantly in WEIRD people, so it does not mean that it is a universal human characteristic: one that has evolved throughout millennia. Even though I agree that the sampling issue is a serious one, and that further experimentation is required to draw more accurate conclusions about altruism in humans, I am slightly more optimistic that we will find this in other cultures. The reason I am more optimistic is because I have provided a coherent culture-gene coevolutionary explanation of how altruistic acts can be selected for. If my theory is true, then I would expect to find it almost all cultures, perhaps only with an exception or two that we can find in all universal facets of the social world. Something that is a hundred percent universal can perhaps only be found in exact sciences such as physics or mathematics.

Furthermore, Stich argues that even if altruistic behaviors were found empirically in all cultures (which would probably take quite some work), then the biologically produced norm-psychology is only part of the picture. He states that “local history, local ecology, and cultural evolution including cultural group selection will be needed to explain which helping norms exist in a given community and why all cultures have norms requiring helping behavior, if indeed they do.” (p. 6). Not all altruistic behaviors are caused in the same way, some might be more culturally learned, or they can be genetically hardwired, like caring for the welfare of one’s children. The latter may have a purely evolutionary explanation, while the former does not: it is mostly shaped by culture. Note that in Stich’s terminology, “a purely

evolutionary explanation” means that it is caused biologically, through the genes. If culture played a role, then there is no purely evolutionary explanation.

Firstly, I don't fully agree with this terminology, because as I have argued before, culture is also an entity that evolves. Stich does seem to acknowledge that there are “cultural forms of evolution including group selection” but does not regard this as a potential evolutionary account of psychological altruism. Thus, the distinction between purely evolutionary accounts and others depending on culture is ill-drawn from my perspective.

Furthermore, the coevolutionary nature of genes and culture implies that we will never have a “purely evolutionary explanation” of psychologically altruistic acts in Stich's account, meaning that they are fully caused by the genes. As we have seen, there are clear cut cases where genes are influenced by culture, for example with lactose intolerance. And we also saw a relationship the other way around, namely with norm-psychology. There are plentiful more examples, and perhaps with future research we will find more ways in which culture and nature influence each other. In any case, I think it is impossible to say that some forms of psychological altruism are *purely* caused by the genes, since it has been in an interactive relationship with culture ever since culture was born.

Conclusion

Biological and psychological altruism have been regarded as something separate in the literature. It was my aim to show that they are not so different after all. Psychological altruism has the extra dimension of motivation attached to it. Our mental activity fuels our motivation, and so it was thought that our thoughts can act on their own, without much influence from our biology.

However, I argued against this by using the Baldwin effect. The main point behind this theory is that culture can have an influence on our genetic makeup. Our physical structure will change to better suit the surrounding cultural environment. This will then further reinforce a particular culture, so that they will both interact and reinforce each other. Such a feedback loop is referred to as the culture-gene coevolutionary theory, which I advocated in this paper.

This co-evolutionary process ensures that altruistic behaviors such as food sharing, which are great activities to ensure the survival of a particular group, are selected for by nature. It was a type of behavior that was reinforced by members of one's group, generation after generation, and thus performed a selection pressure for certain genes to come to expression. Thus, culture performs pressure together with nature to produce successful behaviors that ensure the survival of a species.

Humans are psychologically wired to learn rules of conduct from their environment. This is our particular norm-psychology argued for by Chudek et al. (2013) and others, that predisposes us for norm acquisition. Empirical data has shown that human infants are particularly well-adapted to learn norms and do well upon others, seemingly from internal motivation without external rewards being required, sometimes even being detrimental. These experiments are quite persuasive, though further experimentation is needed to become more certain that this is a universal human phenomenon, rather than being a culturally specific learned behavior.

Thus, the altruistic behaviors that humans portray, which appear to be detrimental to our survival in the evolutionary sense, are not so detrimental after all. Our bodies are constructed in a certain way that makes it pleasurable and feel nice to help other individuals. The evolutionary benefit is not always immediately obvious, like when sharing food with a pigeon at the train station. But even then, this behavior can be explained by a general food sharing trait, that has been selected for with a culture-gene coevolution. Other times, the evolutionary benefit is pretty obvious, like when helping next of kin. This ensures that 50% of one's genes can survive and continue to reproduce.

So, psychologically altruistic acts are motivated by a general human psychological predisposition to be helpful to members of our group. We are selected for to derive pleasure from helping others. This has developed so far, and has become so reinforced that it has become the new standard. It has become human nature to be helpful.

List of References

- Ananth, M. (2005). Psychological altruism vs. biological altruism: Narrowing the gap with the Baldwin effect. *Acta Biotheoretica*, 53(3), 217-239.
- Baldwin, J. M. (1896). A new factor in evolution. *The American naturalist*, 30(354), 441-451.
- Batson, C. D., Van Lange, P. A., Ahmad, N., & Lishner, D. A. (2003). Altruism and helping behavior. *The Sage handbook of social psychology*, 1, 279-295.
- Batson, C. D. (2014). *The altruism question: Toward a social-psychological answer*. Psychology Press.
- Blows, M. W., & Schwarz, M. P. (1991). Spatial distribution of a primitively social bee: Does genetic population structure facilitate altruism?. *Evolution*, 45(3), 680-693.
- Boyd, R., & Richerson, P. J. (2006). Solving the puzzle of human cooperation. *Evolution and culture*, 105-132.
- Chudek, M., Zhao, W., & Henrich, J. (2013). Culture-gene coevolution, large-scale cooperation and the shaping of human social psychology. *Signaling, commitment, and emotion*, 425- 457.
- Cialdini, R. B., Baumann, D. J., & Kenrick, D. T. (1981). Insights from sadness: A three-step model of the development of altruism as hedonism. *Developmental review*, 1(3), 207-223.
- Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection. Or the Preservation of Favoured Races in the Struggle for Life*. H. Milford; Oxford University Press.
- Dawkins, R. (1976). *The Selfish Gene* (1st ed.). Oxford University Press.
- Dennett, D. C. (2017). *From bacteria to Bach and back: The evolution of minds*. WW Norton & Company.

- Doris, John, Stich, Stephen, Walmsley, & Lachlan (2020). Empirical Approaches to Altruism. *The Stanford Encyclopedia of Philosophy*. Retrieved from: <https://plato.stanford.edu/archives/spr2020/entries/altruism-empirical>.
- Dunfield, K., Kuhlmeier, V. A., O'Connell, L., & Kelley, E. (2011). Examining the diversity of prosocial behavior: Helping, sharing, and comforting in infancy. *Infancy*, 16(3), 227- 247.
- Gomes, C. M., Mundry, R., & Boesch, C. (2009). Long-term reciprocation of grooming in wild West African chimpanzees. *Proceedings of the Royal Society: biological sciences*, 276, 699-706.
- Hamilton, W. D. (1964). The genetical evolution of social behaviour. *Journal of theoretical biology*, 7(1), 17-52.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and brain sciences*, 33(2-3), 61-83.
- Koonin, E. V., & Wolf, Y. I. (2009). Is evolution Darwinian or/and Lamarckian?. *Biology direct*, 4(1), 42.
- Laland, K. N. (2008). Exploring gene–culture interactions: insights from handedness, sexual selection and niche-construction case studies. *Philosophical transactions of the Royal Society B: Biological sciences*, 363(1509), 3577-3589.
- Marshall, J. A. (2011). Group selection and kin selection: Formally equivalent approaches. *Trends in ecology & evolution*, 26(7), 325-332.
- Mullen, E. K., & Thompson, G. J. (2015). Understanding honey bee worker self-sacrifice: A conceptual–empirical framework. *Advances in insect physiology* (Vol. 48, pp. 325-354). Academic Press.
- Newton-Fisher, N. E., & Lee, P. C. (2011). Grooming reciprocity in wild male chimpanzees. *Animal behaviour*, 81(2), 439-446.
- Nowak, M. A., Tarnita, C. E., & Wilson, E. O. (2010). The evolution of eusociality. *Nature*, 466, 1057-1062.
- Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental psychology*, 44(3), 875.
- 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.
- Smith, J. M. (1964). Group selection and kin selection. *Nature*, 201, 1145-1147.

- Sasaki, J. Y. (2013). Promise and challenges surrounding culture–gene coevolution and gene–culture interactions. *Psychological inquiry*, 24(1), 64-70.
- Sober, E. (1993). Evolutionary altruism, psychological egoism, and morality: Disentangling the phenotypes. *Evolutionary ethics*, 199-216.
- Sober, E., & Wilson, D. S. (1999). *Unto others: The evolution and psychology of unselfish behavior*. Harvard University Press.
- Stich, S. (2016). Why there might not be an evolutionary explanation for psychological altruism. *Studies in history and philosophy of science part C: Studies in history and philosophy of biological and biomedical sciences*, 56, 3-6.
- Tabibnia, G., Satpute, A. B., & Lieberman, M. D. (2008). The sunny side of fairness: Preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science*, 19(4), 339-347.
- Travisano, M., & Velicer, G. J. (2004). Strategies of microbial cheater control. *Trends in microbiology*, 12(2), 72-78.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The quarterly review of biology*, 46(1), 35-57.
- Turney, P., Whitley, D., & Anderson, R. W. (1996). Evolution, learning, and instinct: 100 years of the Baldwin effect. *Evolutionary computation*, 4(3), 4-8.
- Warneken, F. (2013). Young children proactively remedy unnoticed accidents. *Cognition*, 126(1), 101-108.
- Warneken, F. (2016). Insights into the biological foundation of human altruistic sentiments. *Current opinion in psychology*, 7, 51-56.
- Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, 11(3), 271-294.
- Warneken, F., & Tomasello, M. (2008). Extrinsic rewards undermine altruistic tendencies in 20-month-olds. *Developmental psychology*, 44(6), 1785.
- Warneken, F., & Tomasello, M. (2013). Parental presence and encouragement do not influence helping in young children. *Infancy*, 18(3), 345-368.
- Weber, B. H., & Depew, D. J. (2003). *Evolution and learning: The Baldwin effect reconsidered*. MIT Press.
- Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature*, 308(5955), 181-184.
- Wilson, D. S., & Sober, E. (1994). Reintroducing group selection to the human behavioral sciences. *Behavioral and brain sciences*, 17(4), 585-608.

