

The Question of Pure Altruism

Tohme Abounader

PSY342A

April 9, 2018

The Question of Pure Altruism

Suppose a stranger throws himself in front of a car to save the life of a child about to be run over or a dying wealthy man decides to give all his wealth to charity at no gains to himself. What motivates that stranger to act, causing absolute harm to himself and death with no reward received in return? What motivates the dying man who has nothing to gain from charity decide to give away his fortunes to complete strangers? One answer is pure altruism, the selfless concern for others' wellbeing at the cost of one's own wellbeing, time, or energy. This would be a satisfactory answer that seems to have been accepted society, that selflessness can truly exist. However, there are serious concerns about the selfless aspect of altruism with some scholars, such as evolutionary psychologists and proponents of the social exchange theory, claiming that it is not simply unlikely for it to exist, but that it is outright impossible, while others counter these claims be it through innate, biologically, or religious conviction of divinely acquired selflessness or Batson's infamous empathy-altruism hypothesis. Regardless of which route one takes, the means to the end do not require proper identification to notice that humans do participate in behavior that penalizes them and benefits others, but the question is can and do they do so independently of self-interest?

Initially we are to ask if such altruistic behavior, or behavior in general, is voluntary. With the understanding that there is a huge contemporary philosophical division between determinism and free will, we must note that if behavior is ultimately deterministic then no act is voluntary and therefore any individuals performing altruistic behavior are acting under compulsion of biology, physics, or divine guidance. Such behavior would then not truly be selfless behavior, but rather the result of manipulation by someone or something that poses harm to or sacrifices this individual to save another. That is, if that is the case then the pure and

benevolent nature of that prosocial behavior is lost and therefore is not truly altruistic. Therefore, we must assume that free will must exist for pure altruism to exist. If free will exists, then actions can indeed be voluntary inasmuch as they are not influenced, nudged, or even the direct result of external or social compulsion. This means that a person should be able to perform prosocial behavior out of his own choice and conviction rather than because such behavior is expected of him or bred into him. This is important because if he is acting because of such external motivations then he cannot be acting without the underlying threat of costs and penalties for inaction that might be more severe than the consequences of taking an action (even death, if we are to accept that reputation and judgment proceeds after death or beyond death). For instance, if external motivations are the motive then the man who jumped in front of the car did so not out of pure selfless altruism, to help that child, but rather because this course of action was primed into him, if not also to avert the social consequences of having to live his life shunned, shamed, or guilt-ridden for inaction in a society that penalizes non-prosocial behavior, especially when it is a question of saving an innocent child, or dying buried with billions of dollars that could have saved many lives.

Therefore, for altruism to exist, it would require that individuals possess free will and act out of their own convictions with little to no external stimuli and threats of sanctions for inaction. The person jumping in front of the car would have to do so because he alone wants to independent of society or biological conditioning. The problem here, of course, is that if the person does so because he wants to, then he is inadvertently benefiting himself. There is absolutely no way around this. This is what the social exchange theory posits.

Others have disagreed with this assessment by presenting a differing definition of altruism. They would contend that external rewards (for reputation) detract from the altruistic

nature of an act while acting to obtain internal rewards (such as good feelings, averting trauma, etc.) in return for prosocial behavior is actually altruistic. The problem with this line of thought is that it requires the convolution of semantics and the changing of the definition of “selfless.” That is, it would require that we not consider seeking internal rewards as an instance of self-interest, which is evidently erroneous since acting for any form of internal or external reward would necessitate that the actor be benefiting from the act and therefore is no longer acting selflessly, or altruistically for that matter. Self-interest is then all-pervasive in human action.

Enter Batson who in his 1991 article in *Psychological Inquiry* titled “Evidence for Altruism: Toward a Pluralism of Prosocial Motives” (Batson, 1991). There Batson argues in favor of an alternative explanation, what he calls the “empathy-altruism hypothesis.” According to Batson, the negation of rewards is not necessary, that altruism can coexist alongside the reception of benefits by the altruist but only inasmuch as “these benefits to self are not the ultimate goal of empathy-induced helping, only unintended consequences” (Batson, 2008). Therefore Batson would seek to look at the root of the act – the motivation for acting beyond the fact that it would benefit the actor. To determine whether the motivation is out of egotistical reasons or altruistic reasons arising from empathy, Batson conducted several studies that attempted to isolate and control for variables of empathy and social consequences. Apparently, he created “high empathy” and “low empathy” situations and compared the participants’ willingness to help someone in need named Carol. These situations were created by simply encouraging the participants to “feel for” Carol or to “remain objective”, respectively. In the “high empathy” situation, the participants were willing to help Carol regardless of external rewards/costs or meeting her again while the “low empathy” situation they were willing to help Carol only when the chance of meeting her was high and when the costs for inaction are high. Of

course this might seem satisfactory on the surface, but the problem with this is that the participants were primed to feel or act in a certain way that to the researchers seemed selfless. However, it needs to be noted that even empathy can be an ultimately self-interested pursuit where the participants felt empathetic towards Carol because of the artificial situational priming by the researchers rather than any innate or benefitless sentiments. That is they would feel bad for not helping Carol and therefore would try to minimize such an undesirable feeling by helping Carol as a result of cognitive manipulation such as the “framing effect” and the priming and not because they legitimately cared for Carol and would render aid to her with no benefits for themselves. The benefits received are subtle. In fact, the participants could even feel ashamed and pressured by the researchers for not offering aid to Carol after being primed for empathy whether it lasts after the experiment or if they consciously realize it or not.

Furthermore as Batson along with Ahmad and Tsang noted in their paper “Four Motives for Community Involvement” (Batson, Ahmad, & Tsang, 2002), they argue that it is not altruism alone that must be looked at for prosocial behavior, but four motives in total: egoism, altruism, collectivism, and principlism. They argue that these motives may all be present at the same time and interplay with a notable addition: “it may not be possible to feel empathy for an abstract social category like *the community, people with AIDS, the elderly, or the homeless*” (Batson, et al., 2002). Continuing, different levels of empathic feelings exist based on four elements, those: “(a) who are friends, kin, or similar to us, (b) to whom we are emotionally attached, (c) for whom we feel responsible, or (d) whose perspective we adopt (Batson, 1991; Stotland, 1969). And, like any emotion, empathic feelings are likely to diminish over time (Batson, 1987, 1991)” (Batson, et al., 2002). In other words, though they believe that selfless altruism through empathy can exist between individuals when sufficiently primed, the extension of such a finding to the

social settings appears to be tenuous at best. After all, emphasizing a focus on caring will spark a need to fulfill ingrained or nurtured sentiments of virtue and charity, that is not surprising and harkens back to “name of the game” experiments influencing selfish vs cooperative tendencies such as the Wall Street Game vs Community Game priming experiment (Lieberman, et al., 2004). Any attempts at arriving at the answer to the question of pure altruism in favor of its existence would necessitate the dissecting the definition, the identification of several factors in interplay, and disregarding any small or indirect benefits received.

The danger here is that even after all is said and done, individuals have an incentive to self-report altruistic behavior in an attempt to engage in virtue signaling either as a result of intentional deception or belief arising out of self-deception (Mijović-Prelec & Prelect, 2010). Even preference falsification whereby individuals conceal their true intentions, beliefs, and later actions can distort true motivations behind actions such as seemingly selfless altruistic behavior by appearing to want to render aid out of sympathy when the underlying motivations are ultimately selfish (Kuran, 1987).

Regardless, though arguments in favor of the existence of true altruism would criticize the circular nature of the selfish altruist claim, they do concede that “There can be no such thing as an ‘altruistic’ act that does not involve some element of self-interest, no such thing, for example, as an altruistic act that does not lead to some degree, no matter how small, of pride or satisfaction” (Burton, 2012) but contend by that same reasoning that the universality of self-benefit arising from human action does not mean that true altruism must be impossible, but rather that these “an act should not be written off as selfish or self-motivated simply because it includes some inevitable element of self-interest. The act can still be counted as altruistic if the ‘selfish’ element is accidental; or, if not accidental, then secondary; or, if neither accidental nor

secondary, then undetermining” (Burton, 2012). For after all, a rich man satisfied with his reputation on his deathbed might still give away all his belongings to charity for nothing in return, because it’s the right thing to do. The point that is then made is that simply because there’s unintentional or a hint of reward for altruistic actions does not necessarily mean that the act itself is self-interested. After all, an individual can find benefits even in the worst and most self-harming situations and actions. It therefore does not follow that the self-harming acts are therefore self-interested because of the miniscule amount of benefit gained.

On the other hand, evolutionary biologists would contend that altruism could not have successfully evolved. Robert Trivers in “the Evolution of Reciprocal Altruism” explains that for altruism to be successful, it would require two basic conditions: memory and high population viscosity (Trivers, 71). A society would have to be small enough to be able to retain memory as a reputational accounting mechanism and repeat interactions with the same people over the course of years, but in large-scale societies this largely fails as memory does not scale properly (though money can act as somewhat of an imperfect measure of social contribution (Kocherlakota, 1996) and population viscosity is largely nonexistent beyond the limited interactions within small social or work groups. Meaning, though scholars like Batson might put forward the notion that true or pure altruism can be observed through empathy in controlled experiments to some degree through priming, it would be extremely difficult to observe such altruism in the real world, especially in a society where direct and indirect interactions between us and millions of individuals occur on a daily basis. The problems here are many, but mainly it’s that self-deception and false signalling are major setbacks for the idea of the existence of true or pure altruism. That is essentially because through the evolutionary process, as best exemplified in by Robert Axelrod’s 1981 “Evolution of Cooperation”, concludes that if a society of pure altruists

encounters a single selfish individual then in non-iterative and iterative interactions (or in Axelrod's case, prisoner's dilemma game theory) will lead to a successful invasion by a selfish individual (represented as "ALL D" in Axelrod's paper). That is to say, during humanity's early development, altruism would have been compromised by a single mutation that leads to a selfish individual and with evolutionary psychology in the mix, then that individual would have increased his survival fitness at the expense of everyone else who, being altruists, are also increasing his survival fitness and sacrificing their own. The selfish individual will be able to maximize his benefits and pass on his (selfish) genes which would lead to a repeated process of exploitation and defection against the altruists. Down the road, as in the repeated simulations of the iterated Prisoner's Dilemma shows, the society of pure altruists and the pure altruists themselves will lose out in the end and potentially become extinct as a result without the ability to successfully pass down their genes. Therefore, the selfish individuals would inadvertently develop selfish society which would harm everyone as a result (lower overall payouts from "ALL D" societies compared to all cooperative interactions (Axelrod, 1981)).

However, such an outcome can be averted as with Trivers and Axelrod through the process of reciprocal altruism or "TIT FOR TAT" strategy where cooperation is the first move whereafter all other moves by the opponent are mirrored in subsequent moves by the participant. This is very important because if a society of "TIT FOR TAT" individuals encounters a selfish individual then they will reciprocate his behavior. Say he refuses to help build the communal storage, then the individuals in that society will refuse to aid him in his own building projects or even deny him access to the communal storage as a result. Even though the people in that society would be acting selfishly they would only be doing so as a response to the other individual's action or inaction only to ensure further cooperation which nets higher rewards. This would

penalize the selfish individual's inaction and it would force him to cooperate for his own self-interest, thus turning the selfish individual into a cooperative self-interested altruist. But this poses a new strategy for the selfish altruist: deception and false signalling. In cases where monitoring is simple, such strategies would prove to be very difficult since people can observe to see if that person's words are being put to action. However where obfuscation results due to large societies and increasing numbers of people, then Dunbar's number dictates that past an approximate number of 200 people will cease treating each other as friends and humans and treat them more as numbers or even objects. That is to say, the ability to monitor others' behaviors becomes very limited and attempts to rely purely on trust take hold. Individuals would find it necessary to rely on trusting that others are doing what they are saying unless proven otherwise. In other words, the selfish individual in a large scale or modern society can pose as an altruist with no accurate means of verification. Even in early human societies that began to increase in size and relationships began to be less consistent, such behavior could have occurred and benefited the selfish individual greatly to such an extent that successful acts of concealed defection and exploitation would have led to an increased ability to provide for himself, his kin, and pass on his deceptive genes through the generations. A skilled deceiver could perform a charitable act or two then capitalize on those acts by turning the spotlight on them, referencing them, all the while adeptly not putting in any more notable instances of cooperation or contribution. The concern thereafter is the realization that if such a process occurred, and it seems likely that it did, then ideals of altruism might be ingrained in our biological programming but not pure altruism, but instead that of selfish altruism or even worse: self-deception where unknowingly to us our acts of alleged pure altruism are the result of our evolutionary biological

imperative to signal our true or false virtue to others in order to achieve reputational gains, avert punishment, and ensure survival.

The process does not end there. If the biological argument does not hold water, then the cultural argument retains the same basic ideas of gene transferral but instead of sexual reproduction it would be liberated through the means of social interactions, child rearing, and manipulation. Where evolutionary psychology necessitates the interaction of genes in a competition for survival, manipulative selfish individuals can still have historically come to dominate modern society by convincing and appealing for people in their society to act altruistically and cooperate when they themselves would defect. This can even be seen in the case of the parent-child relationship where the parent stands gain immensely from teaching the child their selected values and ensuring that the child can adequately benefit the parent when the latter becomes unable to provide for themselves. A child, in a sense, would be a parent's insurance policy for later years. A selfish individual, if he manages to convince society to act altruistically, can exploit such a system to his benefit, and if such a society continues down the road with a culture of selflessness and altruistic behavior then selfish individuals can also continue down the same road of manipulation and exploitation without the need for any gene selection. Charity then benefits the selfish individuals who are not ashamed of abusing it. And this gain does not have to be in the form of money, but in the form of respect, reputation, and favor that would be much more powerful than money alone. In primitive tribal societies, popularity would be extremely important in maintaining one's status and value for the tribe so as to not be left out for the wolves over someone else. Even in today's society, reputational gains and popularity are extraordinarily important that could pave the way for social benefits and

facilitation that would have otherwise been extremely unlikely to achieve (even for something as lowly as getting a free entrance to the dining hall because the cashiers like you).

Nevertheless, a 2014 study concludes that given the ability to mitigate suffering, people would choose, on average, to “sacrifice more money to reduce a stranger’s pain than their own pain” (Crockett, Kurth-Nelson, et al., 2014). In a 2014 interview with lead author Molly Crockett noted her expressing the following, “Although I’m fairly confident that the volunteers in our recent study were not making altruistic choices out of concern for their reputation, we cannot rule out the possibility that they behaved altruistically in order to avoid feeling guilty, or to feel good about themselves, rather than because they truly cared about the suffering of others.” Crockett then goes on to ask whether figuring out if true altruism exists even matters, quoting a Stanford neuroscientist Jamil Zaki, and then explaining that unlocking the mechanisms behind altruism and increasing it dwarfs asking if true altruism really exists (Crockett, 2014). Conversely, in a 2011 meta study of 129 dictator game studies by Christoph Engel where one participant receives a sum of money and can then voluntarily impart them money to another participant concludes, “While 63.89% violate the income maximisation hypothesis, 36.11% do not. Hence more than a third of a typical population does indeed consist of subjects who have no reticence to leave a recipient with nothing, although this recipient is at their mercy. Clearly, generosity is not a human universal. Moreover, those who are willing to make a donation do neither give everything, nor do they split the pie equally. On average, they give 42.64%. Hence even those who in principle are generous to a degree exploit the opportunity to their advantage. Even generous subjects thus tend to have a selfish side.” Given what was found by Axelrod regarding the successful invasion capability of an “ALL D” individual in a cooperative altruistic

society, this finding is indeed frightening especially as our societies grow ever larger and more independent with less reliance on cooperation between individuals (Malcolm, 2014).

Regardless of the history, development, and current consequences of altruism, one thing can be known for sure. In the case of the rich philanthropist on his deathbed giving away all his wealth to charity, it may seem like an ultimate selfless act, but it too may be as a result of self-interest and satisfaction. After all, the perseverance of memory is certainly one way through which a form of immortality can be achieved just as reputation persists in the world after the person dies, if not in the potential afterlife where worldly actions would be judged by an ethereal being for goodness. Either way, such an act would still be susceptible to selfish gains, especially if physical and genetic survival is not the ultimate goal for humans but rather their perseverance in memory and reputation. That might be why the man jumped in front of the car to save the child, not because he's selfless, but because he wanted to do the right thing, die doing something that he believes is in line with his beliefs, and be remembered as someone who sacrificed themselves (highest cost) to save a child's life. Consciously or unconsciously he succumbed to the spur of the moment and gambled for something of potentially very high value.

References

- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *science*, 211(4489), 1390-1396.
- Batson, C., & Shaw, L. (1991). Evidence for Altruism: Toward a Pluralism of Prosocial Motives. *Psychological Inquiry*, 2(2), 107-122. Retrieved from www.jstor.org/stable/1449242
- Batson, C. D. (2010). Empathy-induced altruistic motivation. *Prosocial motives, emotions, and behavior: The better angels of our nature*, 15-34.
<http://portal.idc.ac.il/en/symposium/herzliyasymposium/documents/dcbatson.pdf>
- Batson, C. D., Ahmad, N., & Tsang, J. A. (2002). Four motives for community involvement. *Journal of Social Issues*, 58(3), 429-445.
http://www.baylorisr.org/wp-content/uploads/tsang_motives.pdf
- Burton, N (2012). Does true altruism exist? *Psychology Today*. Retrieved from www.psychologytoday.com/us/blog/hide-and-see/201203/does-true-altruism-exist
- Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences*, 111(48), 17320-17325.
- Kocherlakota, N. R. (1998). Money is memory. *journal of economic theory*, 81(2), 232-251. <https://www.minneapolisfed.org/research/sr/sr218.pdf>.
- Kuran, T. (1987). Preference falsification, policy continuity and collective conservatism. *The Economic Journal*, 97(387), 642-665.
https://www.jstor.org/stable/2232928?seq=1#page_scan_tab_contents
- Kuran, T. (1997). *Private truths, public lies: The social consequences of preference falsification*. Harvard University Press.

Liberman, V., Samuels, S. M., & Ross, L. (2004). The name of the game: Predictive power of reputations versus situational labels in determining prisoner's dilemma game moves.

Personality and social psychology bulletin, 30(9), 1175-1185.

<http://journals.sagepub.com/doi/abs/10.1177/0146167204264004>

Malcolm, L. (2014, August 19). Research says young people today are more narcissistic than ever. Retrieved April 11, 2018, from

www.abc.net.au/radionational/programs/allinthemind/young-people-today-are-more-narcissistic-than-ever/5457236

Mijović-Prelec, D., & Prelec, D. (2010). Self-deception as self-signalling: a model and experimental evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1538), 227-240. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2827460/>

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly review of biology*, 46(1), 35-57.

<http://greatergood.berkeley.edu/images/uploads/Trivers-EvolutionReciprocalAltruism.pdf>