

**Evolutionary Explanations of *the Trolley Problem*: Evolutionary Origins of Human
Morality**

Anya Sager

Honors Department: State University of New York at New Paltz

Glenn Geher

21 December, 2022

Abstract

The Trolley Problem was originally described by philosopher Phillipa Foot (1967). The problem starts with a runaway train that could go one of two ways; if you (the operator) do nothing, then the trolley will kill five people (track A), but if you switch the tracks (track B) it would kill one person. There has been further research about the individual used for the action based track that only would kill one civilian. Past research using this paradigm has examined various factors, such as the age of potential victims and the relationship to the operator. From an evolutionary perspective, advancing one's genes into the future is something of a bottom line. This can happen directly, through reproduction, or indirectly via helping kin. Past studies have shown that various factors come into consideration when choosing track A or B: age, gender of the person on the track and the participant, genetic relatedness, and relationship status. The evolutionary moral perspective provides a powerful framework for examining all the different factors that affect these decisions within one model.

What is Evolutionary Psychology

Evolutionary Psychology is a branch of psychology that is derived from biology; presently, it is a largely growing field that has gained a lot of respect due to its scientific nature. Charles Darwin (1859) can be coined as the founder of evolutionary psychology; he founded the principles that evolutionary psychology is rooted in, he also predicted that his findings would change the foundation of psychology (Buss & Shackelford, 1997). Darwin explained that there are struggles in humankind, for existence and mates as examples, but evolutionary psychology goes on to explain why and how these struggles affect our behavior. According to Geher (2014), evolutionary psychology explains human behaviors and psychological processes as products of evolutionary forces such as natural selection through an evolutionary lens. Evolutionary psychology has shown that our minds have evolved. Since our minds have been shaped by evolutionary forces such as natural selection, we can understand all kinds of human behaviors using an evolutionary lens.

One of the largest known evolutionary concepts is altruism. Whether you have heard it in an academic setting or in Taylor Swift's latest album, altruism is a common phenomenon that all humans engage in. Behavior that benefits another organism that is not kin and has an unfit consequence on the individual engaging in said behavior (De'Jesús, 2021). To put it in the simplest of terms, altruism is helping someone else, often at a cost to oneself. One form of altruism is kin selection altruism. This form of altruism is shown by investment of offspring and co-operation of siblings to promote survival and reproductive ease with anyone that may share some genetic codes (Thompson & Fitzgerald, 2017).

Another form of altruism is reciprocal altruism. According to Thompson and Fitzgerald (2017) altruism is not always directly related to kin but rather to an emotional connection.

Through reciprocal altruism we each have made a community for ourselves in a largely populated society but it no longer is kin-based (Garvey, 2018). We also are only altruistic to those who we think will be altruistic back to us. Garvey (2018) uses ancestral hunting as an example. When humans did not have a reliable food source, a few groups of men would go out hunting together and collaborate on finding a kill for meat to feed their community. Every few days, someone would get a kill. This means that one community would get meat for a meal that night since only one person actually killed the animal. However, if that one hunter thought another hunter would share their kill in order to feed both communities (or they had actually done so in the past) then the hunter with a kill would share with the other community, because they know it will help them in the future. We act this way consistently, doing favors for friends when they are going through a hard time, driving your child's friend home today so you do not have to pick them up tomorrow, etc. Altruism ancestrally came from kin and wanting your genes to survive, and it has now become a way to live an easier life.

Evolutionary Moral Psychology

From an evolutionary perspective, human morality evolved under small-scale, ancestral conditions (see Geher, 2014). In ancestral times, any type of betrayal led to excommunication from your close-knit 150 kin related community (Geher, 2014). Usually after a betrayal is committed, the betrayer is called out by other group members, causing a form of guilt, which allows the betrayer to avoid being sent out of the community by showing they feel bad for their actions. The perceived idea of feeling guilt also helps one determine if an action is morally right or wrong (Garvey, 2018). Due to the historically severe responses to betrayal, humans often subconsciously think about the guilt they would feel before making a decision, which then becomes a factor in the decision of betraying someone (Garvey, 2018). Moral emotions and

actions, such as guilt, forgiveness, and apologetic behaviors seem to have all evolved to help people stay in the good graces of others.

Moral Attitudes stem from cognitive processes that have been adapted by experiences such as social interactions, these experiences currently and therefore our cognitive processes currently have not changed much from how they were experienced historically (Garvey, 2018). Some moral decisions and judgements happen subconsciously, feelings like approval or disapproval and avoiding doing things that make you feel bad, come and go like the feeling of hunger (Garvey, 2018). You do not try to feel that way nor do you pay attention to why it is happening, you just have these feelings, but there is evolutionary moral reasoning happening behind the scenes of those feelings and decisions. One reason we know these moral attitudes have evolved over time is because cross-culturally there is evidence that morals are consistent between cultures that evolved never interacting with one another (Garvey, 2018). This means that evolutionary moral psychology can be supported because we know that modern cultures are not interacting with other evolved cultures, they all developed on their own with similar morals.

The Trolley Problem and Our Evolved Moral Psychology

Research on our evolved moral psychology has taken various forms over the years (see DeJesus et al., 2021). Recently, the renowned “trolley problem” has been used by some evolutionary psychologists to ask evolutionarily informed questions about moral behavior. In the Trolley Problem, the operator making the decision about which track to choose, subconsciously thinks about who is on each track and what would cause the least betrayal, leading to the least amount of guilt (Garvey, 2018).

The Trolley Problem was created by philosopher Phillipa Foot in 1967. It is a well-known scenario in which a person is operating a trolley: there are two tracks ahead of them

and the brake is broken. Based on the original use of this methodological paradigm, on track A there are five people working the tracks and there is no hope of them escaping. On track B there is one person working on the tracks, again with no hope of escaping. Track A is the track uninterrupted in front of the operator, meaning if they touch nothing on the trolley, those five people will die. Track B is a track that the operator would have to turn onto, choosing to kill that one person.

This may seem like a fairly easy dilemma, why kill five people when there is an opportunity to only kill one. Philosophers and most others would agree that turning the trolley to track B is the only moral option. Thomson (1985) puts an interesting perspective on the trolley problem that may change one's mind. He says, imagine you are in the scenario of a surgeon. You have five dying patients, but one healthy patient could be killed and his organs could be used to save all five patients. You tell this healthy patient the situation and they say no. Is it still okay to operate on the one person? Most would say no, killing one is not worth saving the five, especially without permission. Though, it is pretty safe to say the worker on Track B does not want to be sacrificed. In reality, you would be saving those five people on track A, not killing them like you would on track B. So here lies the dilemma; would you still switch the trolley to track B to save the five workers?

What Evolutionary Psychologists Have Learned about the Trolley Problem

Now that the Trolley Problem is understood, let us view this from an evolutionary standpoint. Why would humans want to save the group of workers over the one? According to Darwin, we as humans evolved to ensure the survival of ourselves but also our group (Thomson, 1985). One would think about how many people their decision impacts, if five people die versus one person dying, the impact of our decision will probably be 5 times greater. For all of these

workers being strangers, it makes sense morally and evolutionarily that the majority of people save the five workers on Track A. Some people will argue that it was the fate of the five people on Track A to be hit. That the operator will have to actively switch the trolley to track B, altering its course. It is more likely that the operator will do nothing and avoid confrontation by changing the course of the Trolley. But, what happens when the 'workers' have faces? In the 1965 Foot version of the Trolley problem, these are ubiquitous strangers, almost gray figures on the tracks, but more recent studies have found that when the person on track B changes from a strange worker to someone who has specified characteristics, or a friend, this trolley problem becomes much more complicated.

In one study (Zhou, Ze Nan, & Zhu, 2017) when a parent was operating the trolley and the person on track B is their child, they always saved their child when Track A also had a stranger on it. But, when it became two strangers on track A, parents chose to kill their child. Guilt plays a big role in these findings. Choosing to end the lives of two people instead of one might lead to the parent feeling more guilt because they know more parents, families and friends will have to mourn when more people are killed. Age is another factor where people choose based on their need to help a relatively larger number of people. Based on the results of one study (Bleske-Rechek, Nelson, Baker, Remiker, & Brandt, 2010), when there was a two-year old on track B, only 34% of people chose to switch to track B. Once you change from the two year old to people from ages 20-70 that drastically increases to 57%-60% of drivers that chose to switch to track B respectively. The younger the Track B person is, the more life they have left to live and ultimately to reproduce.

Kin in general also are saved more often. When there is no relation to the person on track B when there are five people on track A, 70% of the people switched to track B, with 12.5%

relation: 51% of people switched to track B, with 25% relation: 46% of people switched to track B, with 50% relation 30% of people switch to track B (Bleske-Rechek, Nelson, Baker, Remiker, & Brandt, 2022). This shows kin-based altruism, the people that can pass on genes that have relation to one's own genes are saved more often. In other words, the less related the track B person is to the person driving the trolley, the more likely they are to be sacrificed for the five people on track A. Although romantic partners are not considered kin, since they are not blood related, they offer great resources for you to pass along your genes. In most scenarios, without a partner it is increasingly difficult to create a line of kin. So, it makes sense that the person driving that trolley would have also chosen not to switch to track B if their significant other is there than when compared to a stranger on track B (Bleske-Rechek, Nelson, Baker, Remiker, & Brandt, 2022). Friends are also saved more often than are strangers. Actually, friends are sometimes saved over distant relatives, showing that chosen relationships can be stronger than a bloodline you have no relationship with (Thompson & Fitzgerald, 2017). This makes sense because someone you are 5% related to that you have never met you would not save over a friend you see on a daily basis. Humans will often take close friends and add them into their family, many people call close friends of their parents aunts and uncles, showing a comparison to a 25% genetic relation.

Unanswered Research Questions

One thing that has not been addressed in current literature is parents being on track B. Although we can infer from the data that Bleske-Rechek, Nelson, Baker, Remiker and Brandt, 2022 recorded. With switching to track B rates being low for older-aged track B candidates and high for 50% genetic relatedness, we can infer that the parental track switch rates would be in the 45 age range with the 50% relation which is 27%. A reason why this may have not been done is

due to ethical reasons, the child could be negatively affected by the study if they have to choose to kill their parents or not, so in the future psychologists have to be careful about the format of the testing. The following Table (Table 1) helps visualize this data (Bleske-Rechek, Nelson, Baker, Remiker & Brandt, 2022):

Table 1: Percent of Times “Operator” Switched to Track B in Relation to % Relatedness and Age

	0% relatedness	12.5%relatedness	25%relatedness	50% relatedness	Average
2 y/o	62	30	28	15	33.75%
20 y/o	75	57	48	48	57%
45 y/o	85	65	55	27	58%
70 y/o	90	50	53	48	60.25%
Average	78%	50.5%	46%	34.5%	

Also, there are a lot of unknowns about how the parent saving a child trial was done. It seems shocking that in this study, only a 2:1 ratio of stranger:child was needed for a parent operator to choose Track B. In this study we are not told whether the participants are parents in real life, which could have a massive effect on the outcome of the study. In a study by Garvey (2018), there were times where a dog was on track B and five people were on track A and people saved their family dog. Also, in all of these studies the people on the tracks are entirely made up. They are gray figures that cannot be differentiated. It would be helpful to see some studies where

virtual reality is utilized. Adding a face and body and what feels to be a real life experience could also drastically change the outcome of some of these studies, especially the parent operator and child on track B scenario.

Bottom Line

Taking all this past research into account, when thinking about our evolved moral psychology, four broad take-away messages seem to emerge. First, Evolutionary Moral Psychology shows that humans try to avoid feeling guilty, which subconsciously influences decision making. The reason we avoid feeling guilty is because usually guilt is related to betrayal and betrayal has bad repercussions, especially when friends and family find out. Second, those who are young are saved more often than those who are older in age. People young in age, babies and children, have more time left to reproduce and they are seen to have so much life left ahead of them that the guilt people feel for ending their life early overpowers the amount of ambiguous people on the other track. Third, those who are more genetically related to the operator are more likely to be saved than those who are less genetically related. People want to pass on their genes, and those that are genetically related to us pass on some of our genes when they reproduce. So, wanting them to survive longer to be able to reproduce leads them to be saved more often. Last, people that the operator knows and are fond of are more likely to be saved than a stranger. The guilt for not saving someone that is in your 150 person community is much larger because you know the people whose lives will be affected by this person's death.

References

- Bleske-Rechek, A., Nelson, L. A., Baker, J. P., Remiker, M. W., & Brandt, S. J. (2010, March 20). *Apa PsycNet*. Journal of Social, Evolutionary, and Cultural Psychology. Retrieved November 16, 2022, from www.jsecjournal.com
- Buss D. M., & Shackelford T. K. (1997). *Human Aggression in Evolutionary Psychological Perspective*. Clinical Psychology Review, Volume 17, Issue 6, Pages 605-619, ISSN 0272-7358, [https://doi.org/10.1016/S0272-7358\(97\)00037-8](https://doi.org/10.1016/S0272-7358(97)00037-8).
- De'Jesús, A. R., Cristo, M., Ruel, M., Kruchowy, D., Geher, G., Nolan, K., Santos, A., Wojszynski, C., Alijaj, N., DeBonis, A., Elyukin, N., Huppert, S., Maurer, E., Spackman, B. C., Villegas, A., Widrick, K., & Zezula, V. (2021). Betrayal, Outrage, Guilt, and Forgiveness: The Four Horsemen of the Human Social-Emotional Experience. *EvoS Journal: The Journal of the Evolutionary Studies Consortium*, 9, 1–13.
- Garvey, Brian. (2018). The evolution of morality and its rollback. *History and Philosophy of the Life Sciences*. 40. [10.1007/s40656-018-0190-5](https://doi.org/10.1007/s40656-018-0190-5).
- Geher, G. (2014). *Evolutionary psychology 101*. Springer Publishing Company.
- Thompson, J. A., & Fitzgerald, C. J. (2017, October 9). *Nepotistic Preferences in a Computerized Trolley Problem*. Current Research in Social Psychology. Retrieved November 16, 2022, from

https://crisp.org.uiowa.edu/sites/crisp.org.uiowa.edu/files/2020-04/crisp_vol_25_7_1.pdf

Wilson, E. O. (2005). Kin Selection as the Key to Altruism: Its Rise and Fall. *Social Research*, 72(1), 159–166.<http://www.jstor.org/stable/40972006>

Zhou, Ze Nan and Zhu, Lei, *The Effects of Parental Care Motives and Time Constraints on Trolley Problem Decision-Making: A Virtual Reality Study*, from <https://ssrn.com/abstract=4258164>