

The big idea Philosophy

The big idea: is it OK to do wrong for the greater good?

Jailed tycoon Sam Bankman-Fried believed altruism justified his actions. Was he right?



Illustration: Elia Barbieri/The Guardian

Kieran Setiya

Mon 3 Jun 2024 12.30 BST

Earlier this year, the cryptocurrency billionaire Sam Bankman-Fried was **sentenced** to 25 years in prison on seven counts of wire fraud. It's safe to say that his life has not gone according to plan. But was the plan itself immoral?

By his own account, Bankman-Fried aimed to accumulate wealth for philanthropic causes: “**earning to give**”, in the idiom of the effective altruist movement, of which he was a supporter. Billboards for his cryptocurrency exchange, FTX, proclaimed: “I’m in on crypto because I want to make the biggest global impact for good.” Effective altruists often advocate utilitarianism, according to which we should promote the greatest net balance of benefits over harms, by any means necessary. In other words, we are justified in harming some - for instance, by wire fraud - if the harms are outweighed by benefits to others. According to the judge who sentenced Bankman-Fried: “He knew it was wrong; he knew it was criminal.” But even if he knew the law, Bankman-Fried may not have believed he was doing anything wrong. After all, he planned to donate billions to help those in need. He may have thought he had the answer to that age-old moral quandary: whether it’s OK to cause harm for the greater good.

Philosophical ideas don’t often go mainstream. Effective altruism is one of them; another is “the trolley problem” - a source of countless social media memes and a prominent plot device in [The Good Place](#). What gets lost in the memes is why the trolley problem matters. The point is not to generate ever more baffling cases of moral uncertainty, but precisely to investigate

Bankman-Fried's quandary.

In the classic case, devised by my late colleague Judy Thomson in 1976, you are a bystander at a switch that will swerve a trolley car from the track it's on - hurtling towards five victims who will surely die when it hits them - to a side track with a single victim who'll be killed instead. Pop culture presentations of it suggest that the issue is knowing what to do: should you flip the switch or not? But the trolley problem starts with the fact that most of us have little doubt: you should turn the trolley to the side track, taking one life to save five. This judgment puts pressure on those who answer "no" to Bankman-Fried's quandary - who believe it's not OK to harm some just because those harms will be outweighed by benefits to others.

■ ■ *Philosophical ideas don't often go mainstream. Effective altruism is one of them; another is the trolley problem*

The question is why, if we are right to flip the switch, is it wrong to push a bystander in front of the speeding trolley, bringing it to a halt? Or for a transplant doctor to kill an innocent patient and use their organs to save five lives - both of which strike most of us as grossly immoral? For decades, ethicists, including Thomson, struggled to reconcile our contrasting judgments when it comes to flipping the switch versus pushing the bystander or murdering the patient: in each case, we take one life to save five. If we can't identify a meaningful moral difference, we ought to conclude that, since it's OK to flip the switch, it's OK to push the bystander or kill the patient after all. This conclusion leads inexorably to a more utilitarian moral view, in which it's all right to cause harm in the service of the greater good. And it permits a moral defence of Bankman-Fried. He may have miscalculated harms and benefits, risks and rewards, but there was a respectable philosophical argument on his side.

The twist in this tale is that Judy Thomson ultimately changed her mind. In an article published in 2008, she questioned the idea that it's right to flip that switch, taking one life to save five. Her argument turns on a variant of the classic case in which you have an additional option: as well as switching the trolley to a track with a single victim, you can swerve it into yourself. Thomson's view is that you're not required to sacrifice your life, but if you don't, you can't then turn the trolley on to someone else, sacrificing them instead. If you wouldn't give your life to save the five, how can you justify the decision to take theirs? This question has force even when self-sacrifice is not an option, as in the case we started with: the absence of an option you wouldn't take should not affect your choice among the options that remain.

The message of Thomson's reversal is that, instead of weakening our reluctance to cause harm, careful reflection on the trolley problem should strengthen it. There is no moral difference between flipping the switch in the original trolley case and pushing a bystander in front of the trolley, not because both actions are right, but because - despite a common intuition - both actions are wrong. We should not flip that switch because we would not, in most cases, be willing to sacrifice ourselves.

There may be situations in which it's OK to cause harm to some for the benefit of others: acting in self-defence, for instance. But it's not permissible to do so whenever the benefits outweigh the harms. We should not allow

some future Bankman-Fried to justify his actions by appealing to the greater good.

Kieran Setiya is a professor of philosophy at the Massachusetts Institute of Technology and the author of [Life Is Hard: How Philosophy Can Help Us Find Our Way](#) (Penguin).

Further reading

[Would You Kill the Fat Man? The Trolley Problem and What Your Answer Tells Us About Right and Wrong](#) by David Edmonds (Princeton, £12.99)

[Justice: What's the Right Thing to Do?](#) by Michael J Sandel (Penguin, £10.95)

[Rights, Restitution, and Risk: Essays in Moral Theory](#) by Judith Jarvis Thomson (Harvard, £37.95)
