

**Paul Katsafanas – *Agency and the Foundation of Ethics: Nietzschean Constitutivism* (Oxford University Press 2013) pp. 267 ISBN: 978-0-19-964507-7.**

Many think that the function of a moral code is to guide our actions, to tell us what is permitted and what is forbidden. Some philosophers go further. They hold that content of morality can be derived from its function. Since morality is supposed to govern action, these philosophers hope we can learn what morality requires of us by investigating the nature of human agency. In particular, they suppose that all human beings have certain aims or goals simply in virtue of being agents, simply in virtue of doing anything at all, and that these *inescapable goals* constrain what we ought to do. For them moral demands are grounded in the pull of these inescapable goals: the goal of morality is the goal of agency.

Most people would agree that there are certain *forms* of action one can't engage unless one has certain goals. For example, I'm not really playing croquet unless winning is one of my goals, unless I am trying to win. True I can put mallet to ball without having that goal (e.g. I'm practising my game or humouring my guests) but then I'm not really playing croquet, I'm just going through the motions. Consequently to really play croquet is to put yourself in a position where you won't be fully successful in what you are doing unless you actually win. Furthermore, if I'm playing croquet in a way that makes it unlikely I'll win, this means I'm playing it *badly*. So from the inescapable goal of croquet we can derive norms that both guide and grade our playing of croquet.

Though life is not a game of croquet, perhaps life or agency is like croquet in that it too has an inescapable goal. You can escape the goal of croquet simply by refusing to play but if agency has a goal, you can't escape that goal so easily; even to kill yourself is to do something, to engage in agency. Suppose for a moment that the inescapable goal of agency is pleasure. Then you don't count as acting at all unless you are at least trying to pursue pleasure and you don't count as acting well unless you are behaving in a way that is likely to be pleasant. We have derived a hedonistic moral code from a claim about the nature of agency.

Any plausible account of morality must do two things. First, it must ensure that morality is inescapable, that moral demands don't fail to apply to you just because you don't feel inclined to follow them. Second, it must make room for the possibility that you fail to follow them, that you can behave badly. Our little argument for hedonism seems to do both. Anyone who acts must have pleasure as a goal and so must be failing if (feeling depressed) they court pain. But the mere fact that they must have pleasure as a goal does not guarantee that they won't in fact court pain, for people often pursue their goals badly.

Paul Katsafanas rejects the idea that pleasure is the inescapable goal of action but he is attracted by the thought that action has an inescapable goal, a goal that grounds morality. He explores a number of proposals as to what that goal might be before settling on the Nietzschean idea that the inescapable goal of agency is the exercise of power, an idea from which he derives a suitably Nietzschean moral code. Katsafanas's thesis is novel and imaginative, both in itself and as a reading of Nietzsche but it is vulnerable to an objection which Katsafanas himself wields very effectively against competing proposals, an objection that prevents us from grounding morality in an inescapable goal of agency, whatever goal we choose.

The problem arises from the fact that the demands of morality are not just inescapable, they are also *weighty* demands that should usually (or perhaps always) override the competing claims of prudence, fashion, social etiquette, law and so forth. Now the mere fact that I must adopt a certain goal does not mean that I must give the goal any great weight. To be playing croquet at all I must be trying to win but I need not be trying very hard. I have other goals here e.g. I might be playing croquet mainly in order to teach someone to play. I'm still trying to win but not very hard because teaching is more important to me winning.

This creates a dilemma for Katsafanas. On the one hand, whatever the goal of morality is, it can't be our only goal. As we've seen people can behave badly precisely because they have other goals that compete with and are not always overridden by the goal of morality. On other hand, one surely can't refuse the

demands of morality just by observing that though one shares the goal of morality (as one must) one doesn't put great weight on it. We are all *obliged* to give great weight to moral goals. But how is this obligation to be derived from the mere fact that these goals are inescapable?

To his credit, Katsafanas confronts the dilemma squarely. He claims that we are all inclined to give great weight to the goal of exercising of power but that doesn't help unless we also have some reason to give in to that tendency. Katsafanas adds that the goal of exercising power is endorsed by the agent upon reflection but we can fully endorse an aim of ours without giving it any great weight. I restrain my competitive instincts in croquet not because I have any problem with trying to win. It is just that winning isn't the only thing in life.