Savulescu

Psychology of Ethics and Towards an Ethics of Shame

In this presentation, I will discuss some recent psychological findings arising from a project I am leading on Resource Stewardship, and then talk about some of the philosophical implications of this project. I would like to thank Nadira Faber, Molly Crockett, Guy Kahane, Andreas Kappes, Lucius Caviola and Jim Everett for doing all the important work.

I will start with one project that looked at the importance of reputation to people's utilization of resources. People are social beings. They care very much about how they are perceived, or their social status. There is very little organized human behaviour which is not to some degree directed towards fulfilling the need for eliciting favourable responses from others.

Reputational concerns are recognized as a prime human motive. Some people have even said that morality is in large part about social signalling. And behaviour in the ordinary world is driven by a concern for a reputation. So we have developed three sub-projects of our work all related to this broad concern: examining the influence of reputation over the consumption of resources and moral behaviour.

The first project is a laboratory experiment, because of time constraints I will pass quickly over the structure of the experiment. But essentially, participants play what's called a Dictator game, a behavioural economics game where participants are given a resource which they can use. And they choose between two lotteries, one which is better for the other partner and one which is better for them. So it's a test of how altruistic or selfish they are. And in one-third of the rounds, the participants choose for themselves. They could either win or lose money for themselves. And two-thirds of the rounds they choose on behalf of a partner, so they could win or lose money for the partner.

The choices are made either in public or in private. There is also a cost to choosing the better lottery for the partner: it requires that they invest a small amount of money, which is always the participant's money, and never the partner's money.

So our research questions were: do participants acts selfishly or pro-socially? Do they choose the better lottery for their partner as often as they choose it for themselves? How much money do participants invest to get a better lottery for their partner as compared to themselves? And more importantly, how do the participant's choices depend on whether those decisions are known publicly or whether they are private?

What we found was that when decisions were public, the participants chose the better lottery for their partner as often as they chose it for themselves. But when their decisions were private, the participants chose the better lottery significantly less often for their partner than they did for themselves. So they were more prepared to act altruistically towards the partner when their decisions were public.

How much money did participants pay on average to get the better lottery? When the decisions were public, they invested as much money for their partner as they did for themselves, roughly a pound. When their decisions were private, they invested significantly less money than for themselves. So in cases where a person's reputation as a helpful person was at stake, participants treated other people equally to themselves, but in cases where their decision did

not impact on their reputation, participants acted more selfishly. So this experiment implies people are aware that not acting selfishly is the right course of action but their interest in helping others is less driven by genuine concern for the outcomes for the other person and more for their own reputation.

This study was important because the costs of helping the partner were small. This was a case of "costly helping" but the cost was small and so it was an "easy rescue" or "easy benefit" to provide for the other. This kind of behaviour is, I believe, a minimal moral behaviour. I believe we have a duty of easy rescue, to provide large benefits to others (or prevent large harms) at small cost to ourselves. This experiment suggests people are more likely to engage in such behaviour when their behaviour is public than private.

The second project was built in collaboration with Paul van Lange, a famous social psychologist from the Netherlands, and was a large-scale field survey in the Netherlands consisting of several thousand people. It looked at different districts within areas of the Netherlands and how much they were willing to donate when asked at the doorstep and correlated this information with various other features such as their household income, religiosity, age, sex and population density. In theory, one might think that the richer people were, the more they would donate, or the more religious they were, the more they would donate, or perhaps that older people might donate more. And one of our hypotheses was that in low-density communities where people are more likely to know each other and have repeated interactions with each other, that lower density communities may be more concerned about reputation.

So population density in this study was taken as a proxy for reputational concern, and indeed we found that what correlated most closely with willingness to donate was not income or religiosity which showed no correlation with willingness to donate but population density, and the lower the density, the more people donated.

The anonymity that exists in high density populations means that reputational concerns are less prevalent. So again reputational concern in this case measured through the proxy of how familiar people were with those within their community, and that correlates with altruistic behaviour.

In the last experiment, reputational concern was addressed in a different way. This was an example of how dangerous reputational concern can be and how policymakers need to be aware that it can lead to unwanted behaviour.

We measured participants' intention to save water that should be for others in a resource dilemma and how their behaviour was influenced by others. Overall participants intended to behave selfishly. They used rather than saved water at their own convenience rather than considering how much will be left for others. But participants whose behaviour was approved of by a friend who simply said, "I think what you are doing is right" became even more selfish. Participants' behaviour whose friends did not go so far as to approve the behaviour but simply said it was understandable (such as, "I think what you're doing is not right but I understand what you're doing") also became more selfish. Even just the mere intelligibility by others of their behaviour led to more selfish behaviour.

So as soon as there is uncertainty about reputation, any confirmation of either a positive effect or no effect on reputation leads to more selfish behaviour. So what are some of the conclusions we can draw? Policymakers often overestimate the power of money and financial motives over people's altruistic behaviour. They tend to underestimate the power of immaterial and instrumental motives such as reputational concerns, but these are very powerful human motivators which I'll address in the second part of the talk in more detail. And these can be harnessed as a way of encouraging more pro-social behaviour.

Another major area of our research looked at default effects. This is a well-known psychological heuristic where people are biased towards the presented status quo. And again, to briefly describe another experiment, in this case we showed that after performing a task when subjects were rewarded with a small amount of money, when they were given a *default* of donating this bonus money to a charity, they donated significantly more to charity when the default was reversed, i.e. the default was keep the money rather than donating it. So simply suggesting to people they might like to donate money to a charity will have a significant effect over their behaviour.

So again this is a mechanism which is being much discussed in the literature for encouraging more altruistic behaviour such as donating organs, but it could also be used in the area of charity.

In the third stream of psychological research that we are doing, this time with Guy Kahane, Molly Crockett and Andreas Kappes, we've been looking at two different factors. Many of our decisions are uncertain in terms of their outcome, but what is *also* uncertain in many cases is *how much* they will affect other people. We distinguish between effect uncertainty and what could be called outcome uncertainty.

When our actions are uncertain (effect uncertainty) it's well-known that people tend to utilise this to behave more selfishly. In this study, we looked at the effect of uncertainty about the impact of actions on other people's well-being as an influence over altruistic or selfish behaviour, and in fact showed that when there was uncertainty about *how* harmful it would be for others, people tended to reduce their unselfish behaviour. So again it's a rather complicated study but we manipulated both the uncertainty over *what* the outcome would be for a partner in a game but also what their partner's baseline level of income was and therefore *how* affected they would be from losing or gaining money and showed quite strikingly that in fact people behaved *more* altruistically when they were uncertain about *how harmful* the particular outcome would be for their partner. Yet when there was uncertainty about the outcome itself, people, as previous studies had showed tended to behave selfishly.

So this is another possibility for devising social policy capable of influencing peoples' behaviour: demonstrating the potentially bad effects on human beings affected by those behaviours. So I want to now talk about how these kinds of research could be used, but before I do that, I want to talk about the phenomenon that I think Dale raised in his talk on shame and performing morally wrong acts.

Shaming has become a part of popular culture through the internet. This is an example of Lindsey Stone, given by Jon Ronson in his recent book on shaming. By all accounts, Lindsay was a very effective care worker, who would play games in her spare time of pretending to smoke in

front of no smoking signs, loitering in front of no loitering signs and so on. She would take photos of this kind of behaviour, and post on Facebook as a joke. But following one of these photos, which was taken in a military cemetery, being publicised around the internet, Lindsay was fired from her job and received numerous death threats. That's an example of the ability of shaming to influence people's lives.

Another example from Jon Ronson's book that demonstrates the growing effect of shaming through the internet is when two delegates at a large tech conference made a joke to each other privately, which I don't really understand, because I'm not a tech geek but apparently it has sexual connotations, and another delegate at the conference overheard this, and tweeted the joke, and informed the conference organisers. Both of these individuals were then fired.

But they subsequently went online explaining that they had been fired, and that they had three children to support and that it was a joke. And in fact the online backlash to their side of the story resulted in the woman herself being fired. In fact, shaming is used by American judges now as an alternative to punishment. Some judges use this kind of punishment in an attempt to bring about rehabilitation of the offender and to prevent overcrowding of jails.

In the UK, the Guardian published the names of a rightwing Nationalist Party, the British National Party, which is regarded as racist, although it is a legal party. One of the members was the chief ballerina for the English National Ballet and the controversy over her being named as the member of the British National party in 2007 eventually led to her retiring later that year.

And more recently, gay activist, Peter Tatchell has threatened to out bishops in the Anglican Church if they don't speak out in favour of gay rights and yet are in fact themselves gay. Apparently there are 20 bishops that are secretly gay.

This technique is also being used along the lines of our earlier discussion of using reputational effects to enhance people's charitable giving. Bob Geldof attempted to use shaming when he reportedly publicly shamed the singer Adele for not contributing to one of his charity concerts. And in a study similar to ours, it was shown that by setting high default donations and making public what previous givers to their friends who are participating in the London Marathon have donated, people are shamed into giving higher levels of donations.

In San Francisco, Marc Benioff has called on his fellow tech entrepreneurs to be more generous to antipoverty programs and has threatened to again, to shame them publicly if they fail to donate. So how should we think about the use of these kinds of heuristics and biases such as reputational concern to bring about certain outcomes?

Many of the implications of this work are uncontroversial. So identifying one's own biases or attempting to balance biases or attempting to use them to enhance control over your own actions are, I think, all relatively uncontroversial. What I want to focus on now is using heuristics and biases to promote more moral behaviour in others.

This is being criticised already in the literature for being manipulative. That is not the area that I want to look at today. I want to focus on the possible application of our research of using reputational concerns to promote more altruistic behaviour along the lines of promoting greater

donation to charity. As Warren Buffett says, it takes 20 years to build a reputation and five minutes to ruin it. If you think about that, you'll do things differently.

So how could we use reputational concern? Well, we could use it to promote donation to charity. We could use it to reduce carbon emissions. We could use it to increase organ donation rates. It's been used in the area of sexual ethics to change our attitudes to sexual practice and we could use it to promote vaccination. The ethical question is, *should* we use shaming and harnessing reputational concerns in these areas?

Now, there is moral disagreement over many things. So for example in the area that I work on, I'm famous or perhaps infamous for advocating a principle of procreative beneficence, the moral obligation to have the best child that one can have within the constraints of opportunity costs. And this implies that it would be wrong to have a disabled child when one could have a healthy child.

And in fact, one of the criticisms that I've received from this is if you claim that it's wrong, if you follow Mill's argument, then blame and some kind of punishment, even if it is social punishment ought to be employed. One woman once described to me the anger she had felt when she took her child with Down Syndrome to a park and a bystander came up to her and said, "why did you have that child, you could have avoided having a child like that". And this is an example where using social disapproval appears to follow from doing what we might believe is the wrong action but that social disapproval is very misplaced.

So what I want to ask is what level of response should we have to moral wrongness? There are many debates around eating meat, gun control, driving large cars, airline travel, drinking and smoking. And on the standard view, if we believe that a behaviour is wrong, we ought to blame that individual. However, I think that our responses should be gradated and we should exercise an ethical choice over how to respond to immoral behaviour. At one end, we might do nothing. For example, in the case of somebody who chooses to have a child with a disability or fails to employ means to avoid that, in my view, there is no place for the use of social disapproval in those cases for a number of reasons.

But as we go up to spectrum to people who, for example, have derogated gay relationships themselves while being gay, some form of social reaction is justified, all the way to, in extreme cases, restricting an individual's freedom. So how should we respond and use public shaming and reputational damage? In my view, this is a punishment that should be reserved and ought not be used even to the extent that it's currently used today and the reason for that is the tribal nature of human beings, the ability to create in-groups and derogate out-groups with little control over the resulting harm. Take the Salem Witch Trials of 1692. It started with some people being accused of being witches due to children having fits and ultimately resulted in 20 executions, a very vivid example of the tribal nature of the human animal.

The consequences of public shaming can be, as we saw through the previous examples, significant: loss of jobs, loss of social networks. In many cases it's often used based on a narrow view of morality and often stifles moral progress as it institutionalises a certain moral vision of what's right and wrong.

So in my view we ought to be moral inclusivists. We ought to include in the moral community people whose views we disagree with, indeed people whose views we think are wrong. We ought to use argument. We ought to use evidence, but not social exclusion. We ought to be tolerant and inclusive of those with moral differences.

The problem with the drift towards social exclusion has been seen through the treatment of homosexual, transsexuals and philanderers. This is a picture of Oscar Wilde who was sent to jail for two years. The judge remarked during sentencing that this was the worst case he had ever tried. He said, “I shall, under the circumstances, be expected to pass the severest sentence that the law allows. In my judgment it is totally inadequate for a case such as this. The sentence of the Court is that each of you be imprisoned and kept to hard labour for two years.” He lived for only three years after he was released, dying at the age of 45. His crime was being a homosexual.

So I’m concerned about the tyranny of morality. I believe we ought to draw a sharper private/public distinction and we ought to also draw a clear line between the demands of morality and this fear of self-interest.

As Bernard Williams said as he argued vehemently against utilitarianism, there is an importance to our own projects and attachments. The presence of an eye makes people behave more generously, and the appearance of three dots has the same effect but reversing the dots does not have the same effect. And this illustrates that it's the perception that an eye creates that somebody is watching us. Maybe it's because I work in the area of practical ethics but increasingly many of our actions are seen to be moral actions.

This can be good and bad. It can be good when the cost to us is minimal, and there is a clear benefit, like eyes encouraging us to contribute the honesty box. Yet it can be deeply invasive when significant parts of our life fall under even moral gaze.

When should we harness reputational concerns? When somebody is acting in a public capacity, when somebody directly harms others and in an illegal or grossly immoral way, when somebody themselves has used social ostracism, when somebody's reputation is undeserved or perhaps when they place information in a public space, but in many cases I believe that we ought to carve out an area of moral obligation either through the law or through moral norms or policy. For example, in the case of increasing the supply of organ donors, I believe we ought to give priority to people who have previously agreed to be organ donors, we ought to move to an opt-out system, and we ought to remove the family veto. But I don't believe that public shaming of people who have not agreed to be on the organ donor registry is an appropriate use of reputational concerns because of the danger that I've described.

So how could we use reputational concerns ethically? They ought to be based on plausibly valuable or right outcomes. But in general, appealing to peoples’ reputation in private with their consent with low-cost opt-out options is preferable.

And it is better to use a carrot not a stick. One final example is Bernadette Cleary, who was given an OBE by Prince Charles. She said, “when I started the Rainbow Trust, there were precious few services around to help families with seriously ill children. My idea was simply to supply an extra pair of hands to help practically and a shoulder to cry on when things became too much. Since it

began, some 12,500 families have been helped. It's extremely gratifying to know that the charity has helped those families during a difficult time in their lives." So this is an example of somebody who was praised and whose behaviour inspired others.

In my view, reputational concerns ought to in general be used positively and not negatively. So one of the criticisms of our project of moral enhancement has been about what the goals of moral enhancement should be. But another important question is how moral should we be and what is the place of morality in our life? And today we risk what I've described as the tyranny of morality. So while I believe that the work that we've done on reputation is extremely important, I think it ought to be used wisely and sparingly. Thank you.