Prolegomena 14 (2) 2015: 193–205

Utilitarianism and Moral Integrity: How Evolutionary Psychology Undermines Our Moral Intuitions

IVAN CEROVAC

University of Trieste – Department of Philosophy, Piazzale Europa 1, 34127 Trieste, Italy
ivan.cerovac@phd.units.it

ORIGINAL SCIENTIFIC ARTICLE – RECEIVED: 14/04/15 ACCEPTED: 03/12/15

ABSTRACT: This paper discuses Bernard Williams' argument according to which utilitarianism is a bad moral theory because, by requiring us to reject conscience and our moral emotions in favour of the "lesser of evils", it violates our moral integrity, itself a deep moral ideal. I discuss the implications of this objection, as well as the answer offered by Peter Railton. He claims that utilitarianism should respect (and not violate or reject) our conscience and moral emotions because, by violating our integrity for the best consequences, we would become demotivated and unable to act in the long run, and would thus decrease overall utility. This paper questions whether Railton's solution adequately answers Williams' objection, and argues that a possible answer should be looked for in recent studies in evolutionary psychology and in the very origin of our moral emotions.

KEYWORDS: Consequentialism, evolutionary psychology, moral integrity, moral intuitions, utilitarianism.

> Hush: How many lives do you think you've cost, how many families you have ruined, by allowing the Joker to live? … And why? Because of your duty? Your sense of justice?
>
> Bruce: I'm no executioner. That's why [our compassion] is so important, it separates us from them. If I kill him I will be crossing the line from which I would never return.
>
> Grayson and Beatty (2006)

Bruce Wayne, a fictional character known also as Batman, has had over the years many chances to kill his nemesis Joker. Bruce could have anticipated that police would not be able to keep Joker imprisoned and that he would

escape, thus killing many innocent people before he could be stopped again. Should Bruce have killed Joker when he had a chance to?

If we deliberate that Bruce should have killed Joker, thus saving the lives of many innocent people, we are following the argumentation typical of utilitarianism, an ethical theory holding that the course of action we should adopt is the one that maximizes the overall happiness (White 2010). Since we can reasonably assume that world w_1, in which Joker is dead and all his future victims are alive, contains greater overall happiness than world w_2, in which Joker is alive and many people are not (because of his killing them), we can conclude that w_1 is morally more desirable than w_2. Thus it seems that our duty is to do our best in order to bring w_1 into existence. Bruce Wayne, however, does not share this idea by arguing that, although he is aware that by not killing Joker he put at risk the lives of many innocent people, he will not kill him because that would "make Bruce one of them" – if Bruce ignores his basic intuitions and commitments, his moral integrity would be violated, and his very identity would come into question.

Bernard Williams was the first to emphasize this line of argumentation – utilitarianism is a bad moral theory because, by requiring from us to reject conscience and our personal ideals in favour of the "lesser of evils", it violates our moral integrity, itself a deep moral ideal (Williams 2004). Bruce does not want to kill Joker (though utilitarians think he should) because killing Joker would violate his moral integrity. In this paper I try to defend utilitarian position by addressing Williams' argument with recent empirical data introduced by evolutionary psychology.

In the first section, I will provide a brief overview of Williams' critique of utilitarianism, emphasizing that this is not a consequentialist argumentation – Williams does not claim that we should reject utilitarianism because the loss of our moral integrity will lead to bad consequences. His approach is completely different and invites us to change our moral perspective – if we think of our moral non-utilitarian feelings and intuitions as indications of what is right or wrong, utilitarian argumentation will have no persuasive power over us. According to Williams, Bruce refuses to kill Joker not because he believes that killing Joker would have the consequence of violating his moral integrity, thus turning him into a monster or making him unable to help other people as he used to till now – but because his very moral integrity prevents him from killing Joker.

Following this interpretation of Williams' thought, in the second section I consider Railton's (1984) defence of consequentialism and claim that it misses the central point of Williams' criticism. Railton argues that consequentialism is not self-defeating by introducing the idea of a sophisticated consequentialist who has a reason to cultivate dispositions (emotions, intui-

tions, and commitments) even if they will sometimes lead him to actions that do not produce the best consequences, because this will preserve his moral integrity and lead to better overall consequences. According to Railton, we can say that Bruce is a utilitarian and still argue that he has utilitarian reasons to preserve and protect his moral integrity, and consequently not to kill Joker. However, it seems that, although Railton position is innovative and intriguing, it approaches the problem presented by Williams from the consequentialist point of view, and in the end does not solve it adequately.

In the final section I argue that, in order to defend consequentialism from Williams' critique, we have to bring into question our emotions, intuitions and commitments, i.e. the very elements that constitute our moral integrity. Once we deny that our emotions and intuitions play an important role in moral epistemology we can no longer think of them "as indications of what is right or wrong", and consequently we can no longer use Williams' argumentation and say that there is no reason why we should follow utilitarian deliberation rather than our moral intuitions. This must be done in an impartial way, not by questioning their effectiveness to produce best consequences (for then we are presupposing consequentialist standards as objective), but instead by questioning the very origin of our emotions and intuitions. If we can, by relying on numerous studies in neuroscience and evolutionary psychology, explain our emotions and intuitions as products of our evolution, we can bring into question their moral force – we can thus reply to Williams without directly imposing a consequentialist perspective. Once we can show that our feelings, at least on certain occasions, are not (or should not be) indications of what is right or wrong, we can start arguing that the initial problem (whether Bruce should kill Joker, thus violating his moral integrity) should be viewed from the consequentialist point of view. Of course, one can always claim that, although our emotions and intuitions may be the product of our evolutionary past, they still play important role in our lives and to disregard them completely would lead to a great loss of utility (or to bad consequences) – this can be a legitimate objection. Only now, but and not before we have brought into question the moral power of our emotions and intuitions, we can use Railton's argumentation to show that we should hold to some of our emotions and intuitions in order to produce the best consequences.

1. Williams' critique of utilitarianism

Why do utilitarians believe that Bruce should have killed Joker? As we've seen earlier, consequentialism attaches value ultimately to states of affairs, and its concern is with state of affairs the world contains or will contain. If we want to know whether the action is morally good or bad, we have to focus on what

comes about if the action is done, and what comes about if it is not. We are thus as responsible for things we allowed or failed to prevent, as we are for the things we've done ourselves. This negative responsibility, Williams suggests, might be viewed as a special application of the so called "principle of impartiality", for it abstracts from the identity of the agent. Principle of impartiality holds that there can be no relevant difference from the moral point of view which consists just in the fact that it benefits or harms accrue to one person rather the another – "It's me" can never be a morally comprehensible reason. Negative responsibility abstracts from the identity of the agent and thus represents the extreme of impartiality. By allowing Joker to live, Bruce is as responsible for the death of thousands innocents killed by Joker as he would be if he himself had killed them. Utilitarianism thus rejects the idea that each of us is responsible only for what he or she does, rather than also for what other people do. There are at least some circumstances in which we are instead also responsible for what other people do, as seems to be shown by Bruce's case.

Few utilitarians who argue that Bruce did the right thing by allowing Joker to live might defend their position by including in the utility calculus other more remote (or less evident) effects, like possible effects on the agent's character or on the public at large. If Bruce decides to kill Joker the psychological effects on him might be bad enough to cancel out the initial utilitarian advantages of that case. However, this seems very unlikely; if one is a rational utilitarian agent, he would not have bad psychological effects (because he feels that he has done a morally wrong thing) as long as he had initially effected a correct calculus. If one is not a thoroughly rational agent, he might have bad feelings, but as those feelings are irrational, they should not have great weight in a utilitarian calculation. In any case, these are only one man's feelings, and their weight must be small when compared with the utility of thousands. Peter Railton, however, endorses this line of argumentation and tries to demonstrate that possible effects on the agent's character might have sufficient influence and change the result of the utilitarian calculus. I'll discuss this approach in the next section. However, it is important to emphasize that the critique presented in this paragraph is not William's main point – he does not base his argument on the idea that it is impossible to preserve moral integrity within the utilitarian framework (though he thinks so); rather, he relies on the idea that utilitarianism, as a purely theoretical moral doctrine, is incompatible with our moral feelings and intuitions.

There is a common utilitarian appeal to the fact that Bruce's refusal to kill Joker would be a kind of "self-indulgent squeamishness" – from the principle of impartiality follows the idea of negative responsibility, and the claim that Bruce should kill Joker. His duty is to act in a way that brings about the best possible consequences, and any excuse that acting in such way violates

his moral integrity is just like saying "It's me" once again. If he fails to do so because of his emotions and intuitions, it seems that he is a morally weak person. Williams (2004: 254–255, italics added) disagrees:

> The most [that the self-indulgent squeamishness appeal] can do, so far as I can see, is to invite one to consider how seriously, and for what reasons, one *feels* that what one is invited to do is (in these circumstances) wrong, and in particular, to consider a question from the utilitarian point of view. When the agent is not seeing the situation from the utilitarian point of view, the appeal cannot force him to do so […]
>
> If he does not see it from the utilitarian point of view, he will not see his resistance to the invitation, and the unpleasant feeling he associates with accepting it, just as disagreeable experiences of his; *they figure rather as emotional expressions of a thought that to accept would be wrong*.
>
> [The self-indulgent squeamishness appeal] essentially tells him to regard his feelings just as unpleasant experiences of his, and he cannot answer the question they pose when they are precisely not so regarded, but are *regarded as indications of what he thinks is right or wrong*.

The main problem is how to convince those who do not share utilitarian point of view to disregard their moral feelings and embrace the utility calculus. As long as people don't disregard their moral feelings, they cannot embrace the utility calculus, for if they do, they will lose their moral identity – morality (in particular, utilitarian morality) will impose itself as something alien and their moral integrity will be threatened. Williams (2004: 255, italics added) writes:

> We are partially at least not utilitarians, and cannot regard our moral feelings merely as objects of utilitarian value. Because *our moral relation to the world is partly given by those feelings*, to come to regard those feelings from a purely utilitarian point of view, as happening outside one's moral itself, is to lose a sense of one's moral identity; to lose, in the most literal way, one's integrity.

To regard our moral feelings merely as objects in utility calculus is to alienate ourselves from our actions and the *source of our actions in our own convictions*. The main problem presented by Williams, it seems to me, in not of practical nature as Railton believes (how to do the utility calculus correctly), but of epistemic nature (what is the basis of our moral knowledge: should we do the utility calculus or follow our moral feelings and intuitions). It seems that Williams (2004) considers our moral intuitions as premises in our moral deliberation: they are not merely feelings we have regarding certain actions and states, but indicators that something is morally right or wrong. If we consider our moral intuitions simply as more or less pleasurable feelings regarding certain actions, and if we consider our moral integrity as a kind of balance and coherence between our moral intuitions and our moral actions, it seems that Railton is successful in proving that utilitarianism is compati-

ble with moral integrity. Williams clearly disagrees, arguing that we cannot regard our feelings merely as objects of utilitarian value. They (at least partly) constitute our moral relation to the world, thus having epistemic value (they are indicators whether an action is right or wrong). This interpretation corresponds well with Williams's (1985: 162) "relativism of distance", the idea that, though we can have a moral discussion on a real confrontation (where a divergent outlook is a real option for us), there can always be rationally irresolvable disagreements, as well as his strong critique of most forms of moral objectivism (and moral universalism), the idea that some system of ethics applies universally. Our moral feelings, Williams argues, are not simply unpleasant experiences, but (at least partly) sources of our moral knowledge, and any moral theory that does not treat them as such (and I believe Railton's sophisticated consequentialism doesn't) violates our moral integrity.

2. Railton's defense of consequentialism

Living up to the demands of morality, Railton (1984, 2003) argues, may bring with it alienation[1] – from one's personal commitments, from one's feelings, from other people, and even from morality itself. He believes that problems about alienation show consequentialist theories to be self-defeating (Railton 1984). Not only that alienation will decrease a level of utility we can have from certain relationships or commitments, thus reducing the overall utility in the world, but it might also make us unable to actively promote the good. Had Bruce killed Joker despite his moral conviction that to kill a person is wrong, the impact on his moral integrity might make him unable (or demotivated) to perform many other actions that would help other people (Railton 1984: 135–136 has a similar example with Juan and Linda[2]). He

[1] Railton (1984: 135) defines alienation as a kind of estrangement, distancing or separateness resulting in some sort of loss. This need not be a loss of something of value, and need not be a bad thing after all. Whether certain instance of alienation is a good or bad thing depends upon what is lost. Thus alienation from some people or cultures can be a good thing.

[2] In Railton's (1984: 150–152) example, Juan and Linda have a commuting marriage and normally get together every other week. One week she seems a bit depressed, so he decides to take an extra trip in order to be with her. If he did not travel, he would save a large sum that could be donated to an international organization that would use it to dig a well in a drought-stricken village. This donation would produce better overall consequences, even if we consider Linda's uninterrupted malaise, Juan's guilt and any ill effect on their relationship. If Juan had had a character that would have led him to perform the better act (act with better consequences, i.e. to donate the money and to dig a well), he would have had to have been less devoted to Linda. It is quite probable that if he were less devoted to Linda, his overall contribution to human well-being would be less in the end, perhaps because he would become more cynical and self-centered. Similarly, if Bruce had no problem with killing villains (i.e. to perform acts with better consequences), he would probably value life less and be less attentive to the others in need, thus producing worse consequences overall in the long turn.

believes that we must demonstrate how our moral integrity can be preserved within the consequentialist framework[3].

Railton begins by drawing an analogy between consequentialism and hedonism. He introduces the paradox of hedonism – if one adopts a hedonist deliberation, one may well prevent himself from having or experiencing certain sorts of relationships and commitments that are among the greatest sources of happiness. This seems to be a pragmatic (not logical) paradox – the hedonist ought not to be a hedonist because the hedonist point of view might prevent him from attaining the fullest possible realization of sought-after values. We can thus distinguish two forms of hedonism: (i) subjective hedonism, according to which one should adopt the hedonist point of view in action, and always attempt to determine which act seems most likely to contribute to one's happiness, and behave accordingly, and (ii) objective hedonism, which states that one should follow that course of action which would in fact most contribute to one's happiness, even when this would involve not adopting the hedonist point of view in action. Objective hedonist might thus observe the actual modes of thought and action of those people who seem most happy (he would find out that they are not subjective hedonists). He may than try do develop in himself the traits of character that seem common in happy lives.

Railton believes that consequentialism faces the same paradox as hedonism, and believes that the paradox can be solved in the similar way. He thus distinguishes two forms of consequentialism: according to (i) subjective consequentialism, one should always attempt to determine which act of those available would most promote the good, and should then try to act accordingly (one follows a consequentialist mode of decision-making), while according to (ii) objective consequentialism the criterion of the rightness of the act is whether it in fact would promote most good (one does the thing that would bring about the best outcomes). Railton embraces objective consequentialism which sets a definitive and distinctive criterion of right action, and it becomes an empirical question which modes of decision-making should be employed and when.[4] Since some goods are attainable only if people have well-developed characters, individuals may more often act rightly if they possess certain enduring motivational patterns, character traits and commitments. Objective (or sophisticated) consequentialist thus has a reason

[3] Railton (1984: 149–150) clearly distinguishes consequentialism and utilitarianism, and rejects the latter by introducing pluralistic approach in which several goods are viewed as intrinsically valuable (happiness, knowledge, beauty, respect, solidarity, autonomy etc.). Though interesting and worthy of further analysis, this distinction will not be further discussed for I believe it has no significant influence on the critique I present at the end of this section.

[4] Someone (like subjective consequentialists) who recommended a particular mode of decision-making regardless of consequences would be a self-contradicting consequentialist).

to include in himself certain dispositions that will sometimes lead him to violate his own criterion of right action. He can believe that an act can stem from the dispositions it would be best to have, and yet be wrong (because it would produce worse consequences than other acts available). One can thus approve of dispositions and commitments to rules that do not merely supplement a commitment to act for the best, but sometimes override it, so that one knowingly does what is contrary to maximizing the good. One can be a utilitarian and still embrace his commitments and moral feelings, as well as act upon them even when they suggest actions that are contrary to maximizing the good. Our moral integrity can thus be preserved, and Williams' critique answered.

Does Railton's defence of consequentialism show that consequentialist theories are not self-defeating? It probably does. Does it, however, successfully answer Williams' criticism? I think it does not. Presumably, Railton is such a firm consequentialist that he tends to deal with an objection to a consequentialist theory as if it would be invariantly directed to the consequences that the theory produces. However, as I emphasized earlier, Williams' critique aims at discussing the very foundations of our moral knowledge, and not the consequences of alienation. On the contrary, Railton (like other utilitarians) regards our moral feelings, intuitions and commitments as mere objects in the consequentialist calculus, and (unlike other utilitarians) only gives them great significance in the calculus. He does not regard them as sources of our moral knowledge.

Although Railton (2003: 154, italics added) writes: "There is a *cognitive element* in affection – it is not a mere feeling that is given for the deliberative self, but rather involves certain characteristic modes of thought and perception", he does not regard it as a possible source of moral knowledge. A sophisticated consequentialist has pragmatic (and not epistemic) reasons to cultivate certain dispositions and commitments. He already knows what is good, and the only question is how to promote it. This is evident in many places in the paper (Railton 1984: 156, 144, italics added):

> Objective consequentialist *sets a definitive and distinctive criterion of right action*, and it becomes an empirical question which modes of decision-making should be employed.
>
> [Sophisticated consequentialist] has a reason to include in himself certain dispositions that will sometimes lead him to violate *his own criterion of right action*.

According to Railton, we already know what is the right action (the action that produces the best consequences), and we use the appropriate dispositions to sustain this action in order to bring it about. Williams, on the contrary, argues that these very dispositions can be viewed as one of the sources of our moral knowledge. I've emphasized earlier that the main problem is

how to convince those who do not share the utilitarian point of view to disregard their moral feelings and embrace the utility calculus. Railton does not answer this question; he seems instead to presuppose that we already share the utilitarian point of view. By disregarding the epistemic role of our moral feelings and intuitions, he does not give the correct account of what our moral integrity is.

I suggest that in order to dispute Williams' claim on his own grounds we have to show that our moral feelings and intuitions, at least on some occasions, should not have the epistemic role Williams has given them. In order to defend utilitarianism (and consequentialism) from Williams' critique we must show that some of our moral feelings and intuitions are in fact irrational and should, perhaps, be treated as important objects in utilitarian calculus, but not as sources of our moral knowledge.

3. The role of moral intuitions in ethics

In traditional moral debates moral intuitions have been considered as foundations of a moral knowledge, and every conflict between the implications of a moral theory and our moral intuitions was considered as a good argument against that very theory. It seems that Williams embraces this standpoint when he argues that utilitarianism is a bad moral theory because it violates our moral integrity. However, this view can be brought into question.

Jonathan Haidt (2001: 814) asked people to respond to a carefully constructed incest story:

> Julie and Mark are brother and sister, traveling together on summer vacation. One night, while they were staying alone in the cabin, they decided that it would be interesting and fun if they tried making love. Julie was already taking birth control pills, but Mark uses a condom too, just to be safe. They both enjoy making love but decide not to do it again. They keep that night as a special secret between them, which makes them feel even closer to each other. What do you think about that, was it OK for them to make love?

Most people are quick to say that what Julie and Mark did was wrong. They try to explain or justify their response by giving reasons like the dangers of inbreeding, psychological distress or the destabilizing effect on the society. However, all these objections are mistaken because the story was constructed to avoid them. Eventually, many people might say something like "I don't know, I can't explain it, I just know it's wrong" (Haidt 2001). The judgments these people reach are based on their intuitive responses and not on the utilitarian reasons they offer. Similar conclusions can be drawn from the researches made on "trolley problems": while most people agree that one *ought* to throw a switch thus saving five people (that otherwise would have been

killed) and killing one person (that otherwise would have survived), when faced with similar situations that involve the use of physical violence they change their mind. Most people think that we *should not* throw a fat person from a footbridge over the track, thus stopping the train and saving five people. There is no relevant numeric difference between these two cases (in both we sacrifice one person in order to save five), yet most people consider throwing a switch as a good action, and throwing a fat man on the track as a morally bad action (Singer 2005). Why?

Recent studies (e.g. Cosmides and Tooby 2008) in evolutionary psychology give us answers to this question, as well as to the question about Julie and Mark incest story. The recognition that our cognitive and moral architecture, and consequently our moral concepts, moral intuitions and moral sentiments are the product of natural selection, and so represent reflections of our evolutionary process, seems difficult to escape. We are naturally determined to have such intuitions because of our evolutionary past; a community that did not condemn incest (when contraception was not available) or physical violence among its members had lower chances to survive. Since communities that did not condemn and discourage these behaviours had to bear additional costs (children with development problems and higher rate of injured members), they were less successful in reproduction through time and were not favoured by selection. Eventually, over the generations, only communities who condemn incest and physical violence among their members survived, and condemnation of these behaviours became intuitive, spontaneous and "common sense" response for the members of these communities (Cosmides and Tooby 2008). Furthermore, it has been shown that the amplitude of costs and benefits associated with particular acts influences the strength of moral disapproval and approval, respectively (Lieberman 2008). Greater cost that certain behaviour imposes upon a community will result with stronger moral intuitions against that type of behaviour. Different answers to trolley problem can thus be explained by appeal to our evolutionary past: faced with millions of years of direct physical violence, we have developed a natural negative reaction towards it. Since it has been possible to harm someone by throwing a switch for just a century or two, we do not have an appropriate emotional reaction to this option. A variation to famous experiment conducted by Stanley Milgram (2009) supports this idea: Milgram was researching our obedience to authority figures by asking participants (who did not know they were test subjects, but thought that they were laboratory assistants) to punish a fake test subject (actually an actor, one of his graduate students) when he made a mistake in a memory test. Apart from showing that people have a strong tendency to obey authority even when it is obvious that they are causing suffering in other persons, Milgram found that people are more willing to execute strong punishments and cause intense pain when they could do it

from a distance, by pressing a button and giving a (fake) test subject a (fake) electro shock. When they had to hold the (fake) test subject's arm physically onto a (fake) shock plate, they were less willing to execute strong punishments and cause intense pain. This clearly shows that we have an inborn negative emotional response towards directly and physically causing pain in others, but such emotional response is much weaker when the same amount of pain is caused from a distance. Consequently, it seems that these intuitions are not reliable nowadays, when we have various contraception methods (in the incest scenario) and can commit a harmful (violent) act without direct physical interference (like by throwing a switch in the trolley case).

We can further support this claim by introducing fMRI imaging, a technology that scans the brain of a test subject and monitors which area is active when a decision is being made. In case of moral decision-making presented in the trolley problem, fMRI imaging has shown that considering more direct and personal options (like pushing a fat person from a footbridge) has been linked with strong activity in the area of the brain that is associated with the emotions, while considering indirect and impersonal options (like throwing a switch) has been linked with activity in other areas of the brain, in particular to those related to cognitive functions (Greene 2002). Furthermore, Greene (2014) has been able to show that those who were ready to push a fat person from a footbridge took more time to make a decision than others – they had a negative emotional response to the idea of pushing someone from a footbridge, but they decided (against their emotions) that pushing a fat person would be a right thing in those circumstances. This shows that, at least in some situations, our emotions or intuitions are not reliable indications of what is right or wrong, but can instead lead us to incoherent or contradictory moral beliefs (e.g. believing that two trolley cases with no relevant difference should be treated differently).

If at least some (and maybe all) of our moral intuitions and commitments have their origin in our evolutionary past, and are not suitable for the moral problems we are facing nowadays, it is difficult to claim that they should still be considered as a source of our moral knowledge. If evolutionary psychology can explain why we have a certain intuition or feeling (e.g. why most people think that incest is morally wrong), we can no longer talk about epistemic power of our intuitions and moral feelings. And if they no longer have epistemic power, they cannot be used as objections to moral theories they are in conflict with. If a moral theory (like utilitarianism) violates our moral integrity, we will no longer be able to say that this is an epistemic problem – we will have a reason to disregard (some of) our moral intuitions and commitments. As I noted earlier, I believe that Williams presents his argument as the epistemic problem for utilitarianism, and it seems to me that the only way to adequately answer the Williams' critique is to confront him on

the epistemic ground. By demonstrating that intuitions and emotions have no epistemic value we undermine his idea that one can refuse the utilitarian perspective simply by introducing intuitions and feelings incompatible with it. I agree that there can be incompatibilities between a moral theory and some intuitions and emotions, but argue that these incompatibilities are morally irrelevant if we can show that corresponding intuitions and feelings have no epistemic value. Williams (2004: 260, italics added) himself stresses that in moral deliberation "*we should not try to regard someone's reactions, impulses and deeply held projects in the face of pattern of utilities, but we should first need to try to understand them*". A good way of trying to understand our reactions and impulses (i.e. our moral feelings and intuitions) is inquiring into their origin, and if that inquiry produces good reasons for questioning their epistemic value (e.g. if evolutionary psychology can show that we are predictably irrational in regard to some moral cases, like the incest story or the trolley problems), we should conclude that (some of) our moral intuitions are not good objection to utilitarian arguments. This, of course, does not entail that our moral intuitions are, by default, epistemically flawed. However, it demonstrates that, in certain cases, we can make a good non-utilitarian argument against moral intuitions[5]. Evolutionary psychology is still developing and more and more morally relevant cases will be analysed – thus evolutionary psychology represents a good non-utilitarian way of evaluating our moral intuitions.

One can claim that, although our emotions and intuitions may be the product of our evolutionary past and have no epistemic power, they still play important role in our lives and to disregard them completely would lead to a great loss of utility (or to bad consequences). Only after bringing into question the epistemic and moral power of our emotions and intuitions), we can use Railton-like argumentations to show that we should hold to some of our emotions and intuitions.[6]

[5] Since Williams (2004: 254) regards our moral intuitions as indicators that something is right or wrong, something important is lost if we can show that our moral intuitions are epistemically unreliable. Furthermore, if moral intuitions constitute an important part of our moral integrity (namely, moral integrity is preserved when our actions correspond to our moral intuitions), then the entire argument relaying on moral integrity is seriously weakened. Consider someone whose moral intuitions are clearly wrong (e.g. a Nazi officer) – it would seem very odd to argue for or against certain action by introducing the importance of the preservation of his or her moral integrity. The problem is reintroduced when it is not so clear that someone's intuitions are right or wrong (e.g. trolley problem), and it is in this cases that evolutionary psychology can give us valuable answers.

[6] I thank Patrizia Pedrini and Elvio Baccarini for numerous discussions and consultations that have inspired and shaped this paper.

Bibliography

Cosmides, L. and Tooby, J. 2008. "Can a General Deontic Logic Capture the Facts of Human Moral Reasoning? How the Mind Interprets Social Exchange Rules and Detects Cheaters", in W. Sinnott-Armstrong (ed.), *Moral Psychology: The Evolution of Morality; Adaptations and Innateness* (Cambridge, MA: The MIT Press), 53–120.

Grayson, D. and Beatty, S. 2006. "Hush Returns", in *Gotham Knights* 74 (Titan Books).

Greene, J. D. 2002. *The Terrible, Horrible, No Good, Very Bad Truth about Morality and What to Do About It* (Unpublished PhD Thesis, Department of Philosophy, Princeton University).

Greene, J. D. 2014. "Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics", *Ethics* 124 (4), 695–726.

Haidt, J. 2001. "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment", *Psychological Review* 108 (4), 814–834.

Lieberman, D. 2008. "Moral Sentiments Relating to Incest: Discerning Adaptations from By-products", in W. Sinnott-Armstrong (ed.), *Moral Psychology: The Evolution of Morality; Adaptations and Innateness* (Cambridge, MA: The MIT Press), 165–190.

Milgram, S. 2009. *Obedience to Authority: An Experimental View* (New York: Harper Collins Publishers).

Railton, P. 2003. *Facts, Values, and Norms: Essays toward a Morality of Consequence* (Cambridge: Cambridge University Press).

Railton, P. 1984. "Alienation, Consequentialism, and the Demands of Morality", *Philosophy and Public Affairs* 13 (2), 134–171.

Singer, P. 2005. "Ethics and Intuitions", *The Journal of Ethics* 9 (3/4), 331–352.

Sinnott-Armstrong, W. 2011. "Consequentialism", in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2011 edition), URL = http://plato.stanford.edu/entries/consequentialism/.

White, M. D. 2010. "Why Doesn't Batman Kill the Joker?", in W. Irwin and D. K. Johnson (ed.), *Introducing Philosophy Through Pop Culture: From Socrates to South Park* (Oxford: Wiley Blackwell), 311–324.

Williams, B. 2004. "Against Utilitarianism", in L. Pojman (ed.), *The Moral Life* (New York: Oxford University Press), 249–261.

Williams, B. 1985. *Ethics and the Limits of Philosophy* (Cambridge: Harvard University Press).