Andrei G. Zavaliy

Saving Morality: A Case against Moral Neutralism

Studia Humana nr 2, 40-52

2012

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

MUZEUM HISTORII POLSKI

Studia Humana
Volume 1:2 (2012), pp.40—52

Saving Morality: A Case against Moral Neutralism

Andrei G. Zavaliy

American University of Kuwait

azavaliy@auk.edu.kw

Abstract:
The purpose of this paper is to defend a position in metaethics, saving morality from certain reductionist attempts, and arguing that a moral point of view denotes a distinct attitude toward the world with a set of relatively stable conditions. I discuss the problem of demarcation between the moral and the non-moral domains, and contrast the two basic approaches – moral neutralism and moral descriptivism. Moral neutralism is defined as a view which builds no content requirements into the definition of moral rules, whereas moral descriptivism or essentialism places identifiable constraints on the content of an action-guiding principle if such a principle is to count as a moral, as opposed to a non-moral, rule. I show that adopting neutralism is tantamount to giving up ethical theory as a scholarly activity with a distinct subject matter altogether. It is further argued that W. Frankena's essentialist definition of morality, as well as a more recent view of Catherine Wilson, share a similar weakness and fall short from neutralizing neutralism. Finally, I propose a modification to the essentialist account of morality, which would significantly increase the resilience of such an account to attempts of reducing moral prescriptions to any action-guiding policy whatsoever, as long as such policy is sincerely adhered to and followed consistently. The proposed modification is described as a *Realism constraint*, and it refers to the connection between one's expectations of the outcome of observing a prohibition or following a certain rule and the actual consequences of following a given policy.

1. The Problem of Demarcation: Neutralism vs. Descriptivism

Drawing a line between the moral and non-moral normative systems is a less familiar procedure than drawing a line between moral and immoral actions. Perhaps, for this reason, it is often assumed that the former distinction is self-evident. Yet, imagining oneself in a position of an anthropologist who discovers a previously unknown exotic culture with a complex system of social relations and a well-developed, albeit a very unusual, code of the socially acceptable behavior patterns, we can readily appreciate the difficulty of sorting out the many rules that the members of that society actually observe or the ones they claim to be important into the moral and non-moral categories. What should we take as the criterion of a *moral* rule? We may call it the *problem of demarcation*[1] between the moral and non-moral spheres.[2] 'The aberration of closeness' generally hides this difficulty when we consider the more familiar social arrangements such as our own society, but the problem is immediately felt as soon as we attempt to give a precise formulation of the distinctive features of the moral rules for conduct as opposed to the non-moral ones.[3]

On the most general level, one may arrive at two different kinds of *substantive* theories of what constitutes the moral point of view proper. To begin with, one may emphasize the close connection between the normative rules that prompt or prohibit certain types of behavior and the actions of agents who claim to accept those rules. On this view, the problem of specifying the *moral* rules (as opposed to the various non-moral guides) is a matter of observing which ones the members of a given society *actually* take as overriding, i.e., as taking unquestionable precedence in cases of conflict with other (non-moral) requirements. Morality, then, becomes coextensive with any action-

guiding code of conduct adopted by an individual or a group of people. And the problem of recognition of moral principles in a society is thus reduced to the problem of identifying those rules, whatever the content, that are actually observed and honored by the majority of members.

I will follow Kai Nielsen in referring to the position just introduced by the name of *neutralism*.[16] Neutralism in its pure form does not build any intrinsic constraints on the form or content of the moral principles, i.e., it remains neutral with respect to what those principles can be about or how they should be formulated.[4] It is rather primarily concerned with the functional, action-guiding character of those rules, and makes the fact of actual observance of the rules by the agents into the *constitutive* part of morality.[5] Jonathan Bennett seems to advocate this position when he writes the following:

> There could be dispute as to whether the springs of someone's actions constitute *morality*. I think, though, that we must admit that someone who acts in ways which conflict grossly with our morality may nevertheless have a morality of his own – *a set of principles of action which he sincerely assents to*,[6] so that for him the problem of acting well or rightly or in obedience to conscience is the problem of conforming to *those* principles. [2, p. 125]

We can further illustrate neutralism as a philosophical position by reference to a relatively recent and influential tradition. J. P. Sartre in his earlier existentialist writings advocates an ethical theory that ties the moral status of an action to a free, unimpeded choice by an individual.[17] There is no natural or supernatural fact, according to Sartre, that should constraint the use of the concept 'moral goodness'. It would be acting in 'bad faith' to justify one's decision, say, to go to war, by reference to some special feature of an action, its foreseeable consequences, an abstract moral principle or concrete advise by another. An action, whatever it is, becomes morally right (and morally relevant) precisely because it is freely chosen and performed by the agent in the absence of any constraints of that kind. Whatever external features of an act are there, it is our free *endorsement* of these features that explains the action, rather than the features themselves. On this view, however, morality ceases to have any definite content, and there can be no meaningful distinction between correct and mistaken use of moral predicates, as long as they are sincerely used. Any type of action or a state of affairs with any set of objective features can become morally relevant through the performative process of endorsement.

One of the consequences of adopting neutralism as a kind of substantive theory of the moral point of view is that we can now legitimately speak of the 'Nazi morality', 'Slave-owner's morality', the 'Ik tribe morality' and the 'Dobu morality'[7] (among others), without being accused of committing a category mistake. However perverse or unjustified were some of the racist principles practiced, for instance, by the Nazis, they were still the kind of principles that the Nazis took seriously and were implementing consistently in their conduct – i.e., the kind of principles that they, using Bennett's phrase, "sincerely assented to." As such, the set of *those* principles would constitute a peculiar *morality* of the Nazis, even if an extremely *bad* morality from our own perspective.[8]

Another implication of neutralism is that certain important categories are eliminated from the ethical discourse. For instance, neutralism, if accepted, would make the notion of *amoralism* all but empty. Indeed, it would deny the possibility of amoralism. An aggressive psychopath whose anti-social behavior can perhaps be subsumed under the principle "Always follow your immediate impulse" or the like, would not be devoid of morality altogether, according to neutralism, but would rather have an unusual moral code of his own. As Richard Garner remarks in criticizing a similar view,

> Someone might say that the mere fact that we choose one course of action over another shows that we have moral principles. When "amoralists" behave as if they subscribe to moral principles, we can conclude that they

do subscribe to moral principles, perhaps without realizing it. [4, p. 283]

The actual behavior of amoralists can always be described as being in accordance with the rules and principles that they take as overriding. The principle of ethical egoism or any other principle that arbitrarily selects the group of people whose welfare is to count for more than the welfare of the outsiders is the kind of principle that can be held "sincerely" and manifested in one's everyday conduct. Those accepting these principles would satisfy the behavioral and dispositional requirements set by neutralism, and can in fact be considered *moralists* (as long as they do not deviate from the kind of behavior sanctioned by the adopted principles) by the people of similar moral convictions.

It is now clear that neutralism in metaethics undermines the foundations of moral philosophy as a distinct discipline. The designation of something as a "moral rule" becomes an honorific rather than a descriptive designation (meaning: "a very important rule"), and the study of moral rules and their influence on human behavior can be safely relegated to behavioral sciences. If *any* rule or *any* natural feature can become a moral rule or a morally relevant feature as long as they function in a certain way, there is clearly no need for a special discipline that would study morality. In this sense, neutralism can be seen as a reductionist theory.[9]

Besides an understandable concern for the future employment of moral philosophers should neutralism succeed, my opposition to the view is driven by a strong intuition shared by many people that, in R. Joyce's words, "a system of values in which there was no place for condemning Nazi actions simply would not count as a *moral* system." [10, p. 43.] The initial plausibility of neutralism depends to a large extent on the plausibility of the assumption that all alternative attempts at specification of the moral point of view that take into account the form and the content of moral principles, and that seek to establish a core set of non-relative moral requirements and objective moral values, fail. If no substantive account of morality can be justified, neutralism appears to be the only option left. Yet, there are reasons to believe that a defensible account of the moral point of view can be given which would encompass the constraints on both the form and the content of the rules of morality.

A *descriptivist* or *essentialist* account of morality implies that certain attitudes, normative judgments, principles, action-guiding systems or particular actions can be *excluded* from the moral realm using the definite criteria set up by the account. Descriptivism may assume many forms, but overall it seeks to restrict the domain of the morally relevant features of the empirical world so that to exclude certain normative systems from the list of *rival* moralities. This account can be characterized as a kind of *moral naturalism*, since it holds that certain factual considerations necessarily count for or against the ascription of moral terms. It further implies that simply using the language of morality and employing moral concepts and predicates in description of one's behavior is not sufficient for taking a moral point of view. Rather, certain objective descriptive criteria[10] need to be satisfied for legitimizing the moral discourse.

In what follows I will look at the two versions of a descriptivist theory, each trying to specify the features necessary for demarcating certain rules or judgments from their non-moral counterparts. One such account was defended by William Frankena, and the other one was recently proposed by Catherine Wilson.[11] Despite their obvious differences, we will observe that both accounts share a similar weakness.

2. The Descriptivist Project

Much has been said and written in defense of the thesis that morality as we know it can be explained by or derived from self-interested considerations of rational agents. Hobbes' classic model suggests a picture of humans, tired of life of unending confrontation, agreeing on certain behavioral constraints in order to improve the chances of their own survival, peaceful existence, personal security and prosperity.[12] This story with changing details was repeated many times by a

number of later philosophers and social scientists. John Stuart Mill, for instance, ultimately argues for the special place of justice in the system of moral values by telling an essentially Hobbesian story in the last chapter of "Utilitarianism". According to Mill, the sentiment of justice derives its intensity from its link to the animalistic need for retaliation. But its moral force comes from the "impressive" kind of utility that is involved in rights violations - namely the interest of security. Justice has value only in so far, and as long as, it contributes to the overall increase of happiness. Furthermore, one can be motivated to promote justice and act morally if it is pointed out that, being the creatures that we are, we shall find the greatest happiness in living lives in accordance with utilitarianism. Mill argues, in other words, that one needs not fear that commitments to impartial requirements of utilitarian morality may incur heavy tolls on one's own well-being. Quite the opposite – as people are educated to become more and more *impartial*, they will see – paradoxically – that their lives are getting better and better *for them*.[14, Ch.3.]

There is an equally authoritative tradition which argues that a self-interested consideration is the paradigm example of non-moral motive, and nothing properly moral can ever appear from such an ignoble source. Kantian ethics is certainly the most well-known representative of this line of thought. On Kant's view, showing that an action was ultimately based on a selfish maxim is sufficient for disqualifying that action from the moral domain, no matter what the consequences for others. Indeed, no action or evaluative judgment can be genuinely moral for Kant unless it is performed on the basis of the maxim which is abstract enough to remain oblivious to the peculiar circumstances of the agent, i.e., his natural inclinations and private interests.[13]

A more recent exposition of the essentially the same intuition is defended by Catherine Wilson in her book "Moral Animals" (2004). Wilson suggests defining morality as *a system of advantage-reducing imperatives*, i.e., as a system of rules and prohibitions which compensate for natural inequalities within a society. She calls it a *semi*-essentialist definition because it points to the proper function of moral rules in addition to the content and form of those rules (e.g., the anonymity requirement). Attempts to provide a functional definition of morality are certainly well-known in the history of ethics. The distinctive function of moral rules and moral virtues was variously seen as contributing to (or being constitutive of) individual happiness, whether here on earth or in the life to come, or pleasing some supernatural being, or else, more influentially, as aiming at creating a possibility for peaceful coexistence of many people with different interests and various capacities within one society. A further development of that latter minimalist description of the purpose of morality may plausibly include its role in increasing the happiness of the greatest number of people, and in minimizing the amount of pain and suffering.

Wilson's own conception of morality is worked out within the context of that classic tradition which sees the essential role of moral restraints in peacefully regulating the relations between several individuals or groups of people. There are, however, important new developments in Wilson's account. Unlike Hobbes, who starts with the premise of universal *equality* of all humans – a condition that he sees as being primarily responsible for the constant war of all against all, and which eventually creates a need for agreeing on general rules regulating behavior, Wilson starts with the assumption of fundamental *inequality* of the members of a social group. It is because people are unequal in several important respects (e.g., physical strength, financial and intellectual power) that we need moral regulation to begin with. More specifically, Wilson sees the primary function of morality in that it controls and inhibits the spontaneous aggressive impulses which otherwise give advantage to the naturally stronger or smarter members of the social group. The goal here is to counterbalance natural inequalities in such a way as to increase the range of possibilities for the weaker members. Hence a paradigm form of a moral rule would be a categorical prohibition on a certain type of behavior, e.g., "Thou shall not do X!" Without this internalized prohibition a stronger individual might see no reason for not taking the property or even life of a weaker member of the group, which in turn would greatly reduce the chances for procreation and survival of the weaker party. The moral prohibitions, according to Wilson, seek first of all to offset the advantage of the naturally favored subjects (i.e., the one who could otherwise do X with impunity), and create a more favorable surroundings for the naturally or contextually disadvantaged ones:

> *Moral rules are restrictive and prohibitory rules whose social function is to counteract the short or long term advantage possessed by a naturally or situationally favoured subject*. A morality, in short, is a system of compensatory or advantage reducing imperatives that correspond to moral judgments.[20, p. 9]

Wilson's view on the nature of morality creates an interesting contrast with Thrasymachus from Plato's "Republic", who saw justice as a normative system that serves the interests of the stronger. On the contrary, Wilson argues that morality, including that part of it concerned with justice, is primarily in the interest of the weaker party. In this she follows in Nietzsche's steps, although without his scornful feelings about the whole moral project.

By emphasizing the preventive, restraining role of moral rules, Wilson captures much of our intuitive thinking about the aim of morality and the content of moral values. We are outraged when we hear that a person was hired or promoted when we also know of his or her close relationships with the boss. The existence of such a relationship, we tend to think, creates an unfair advantage, which excludes the equally qualified candidates from consideration from the start. The morally right thing to do in the present context, we might agree, is to create a more equal opportunities for all parties and to exclude favoritism.

Still, Wilson's characterization of morality seems to be both too broad and too narrow, as she herself recognizes. She considers the following objection: "Many moral rules are not compensatory or advantage-reducing, and many advantage-reducing or compensatory rules are not moral." [20, p. 13] The first category would include some highly esteemed prescriptions (e.g., a commandment to rest at a certain day of the week) which are hard to interpret as being in the interest of a disadvantaged party. The second category might include ordinary rules of prudence, certain game rules, as well as, say, laws that exempt poor families from paying taxes. All this amounts to saying that Wilson's biggest challenge comes from the side of those theorists who deny that there can be any "cross-culturally valid characterization of a moral rule" and who argue that "to designate a rule as moral is [merely] to single it out as an especially important personal conduct rule," [20, p. 14] i.e., from a position that we have earlier identified as neutralism.

Wilson attempts to meet the neutralist's challenge by offering two responses. First of all, a semi-essentialist, when faced with the *prima facie* non-compensatory rules which are nonetheless singled out by a culture or a group of people as especially important rules of personal conduct (e.g., rules prescribing a certain dress code) may simply insist that those who assert these kind of rules "are speaking or writing in an unusual dialect. [These rules] do not constitute moral rules; they are assignable to the neighboring category of restraining usage taboos applying to objects belonging to a sovereign entity, oneself or perhaps God."[20, p. 15] This initial response, however, amounts to a mere denial of the objector's view on the nature of morality, without yet giving an independent argument showing what is wrong with the neutralist's position.

But such an argument is clearly needed. Wilson's historical reference to the moral codes of antiquity can be seen as an attempt at providing an independent support for the general idea behind her interpretation of the moral rules. She cites historical evidence showing that much of what passed for morality in Ancient Egypt in the 3rd millennium B.C.E. had to do with the rules prohibiting taking advantage of the weaker subjects, such as orphans, widows, poor strangers, etc. It is not clear, however, how much does this single reference to Ancient Egypt help to substantiate the case. One may argue that historical evidence can support a certain view on the constitutive features of morality only if it is representative of all or most known cultures. In this case, however, one feels that the example of funerary inscription from the Egyptian Old Kingdom was tendentiously chosen, and that equally ancient moral norms (e.g., from the Old Testament) were unfairly ignored. But the latter contains many prohibitions which appear to have moral authority, and yet can hardly be interpreted as serving to offset the advantage of the stronger. By analogy, one would rightly feel suspicious about an argument for the truth of the Divine Command Theory which is supported by reference to the moral justificatory practices of the Old Testament.

Wilson's other response to the neutralist's challenge is somewhat more evasive. The strategy here is to concede that many of the presented action-guiding rules that, on the first glance, do not seem to function as advantage-reducing strategies (e.g., duties to self, rules prohibiting cloning or recreational drug-taking) have yet "the proposed moral marker to some degree" [20, p. 15] on the deeper analysis. A policy that, for example, prohibits the vending of surgically extracted organs may ultimately reflect worries about the "temptation to victimize helpless or needy persons that the profitable vending of organs would awaken;" and the prohibition on recreational drugs may be seen as a moral issue only if "such persons are conceived as escaping ordinary demands and responsibilities and as letting others down." [20, p. 16] Wilson admits further that these 'worries' and 'conceptions' "may be foolish worries or not," [20, p. 16] i.e., a certain rule may acquire the moral marker and thus become a moral rule in so far as the background justificatory story refers to its alleged advantage-reducing effect, but it need not be *true* that the rule in question in fact has that effect. Hence, Wilson concludes: "The semi-essentialist need have no objection to including rules mandating, say, women's hair covering, as moral rules, provided they are not taken without further explanation to be examples of central or focal moral rules and provided their advantage-reducing feature can be made apparent." [20, p. 16]

Wilson's hypothesis enters dangerous grounds when she concedes to the critic that a rule can become a moral rule depending on the background justification. The one and the same rule, e.g., 'Share your candy with your friends' "may function as a moral injunction not to tolerate the relative deprivation of others, or to be a prudential recommendation for achieving popularity." [20, p. 14] The claim that a social rule can have dual significance, as both moral and non-moral prescription, would not be accepted by some moral theorists (e.g., utilitarians), but is a familiar point in other moral traditions (e.g., Kantianism). There is, however, one important question that the defenders of the 'dual significance' theory must address. It is the question of the criterion that one should use in deciding whether a rule in question is moral or not. Wilson cannot mean that the problem of demarcation can be settled by looking whether the behavior sanctioned by the rule would indeed have a compensatory effect, since it is the *same* rule (and thus the same behavior with the same consequences that is sanctioned by the rule) that can be both moral and non-moral. So the demarcation must depend on some subjective features of the agent (e.g., his background beliefs or intentions) who adopts or prescribes the rule in question. But then, one might argue, Wilson's (semi)essentialist account of morality collapses into neutralism, since now practically any rule formulated in a certain way (so as to respect, for instance, the anonymity requirement) can be taken as a moral rule given the right background beliefs of the agent. "Everyone must sacrifice their firstborns to the god of the great river" may well be interpreted as a moral rule when combined with a sincere belief of a local shaman that otherwise the angry god will destroy the crops of poor villagers. I take this consideration as being quite crucial for the survival of moral essentialism.

We can somewhat clarify the objection just mentioned by considering another attempt to escape the claws of neutralism. In an account that has approached a classical status by now, William Frankena offers a plausible characterization of morality and it will be helpful to quote his description of the moral point of view at some length here:

> My own position, then, is that one is taking the moral point of view if and only if (a) one is making normative judgments about actions, desires, dispositions, intentions, motives, persons, or traits of character; (b) one is willing to universalize one's judgments; (c) one reasons for one's judgments consists of facts about what the things judged do to the lives of sentient beings in terms of promoting or distributing non-moral good and evil; and (d) when the judgment is about oneself or one's own actions, one's reasons include such facts about what one's own actions and dispositions do to the lives of other sentient beings as such, if others are affected. One has a morality or moral action-guide only if and insofar as one makes normative judgments from this point of view and is guided by them. [3, pp. 113-14.]

Frankena lists here four distinct characteristics of the moral domain. The very last sentence in the quote above suggests that he takes each of these four features to be necessary and jointly sufficient conditions for taking the moral point of view.[14] Arguably, accepting such a precise specification of the moral domain will not yet commit one to any particular theory in normative ethics. It is likely that a number of different and, perhaps, conflicting ethical theories would satisfy the four requirements. This fact will be relevant in the sense that it will tend to confirm that we are dealing here with the theories of *morality*, but it will yield no additional support for any of the specific accounts of the nature of moral obligation and moral rightness.

The initial feature deals with the possible objects of evaluative judgments. Frankena lists actions, desires, dispositions, intentions, motives, persons, or traits of character as the only acceptable objects of a *moral* normative pronouncement. First, the most basic constraint here is the constraint on the domain of the possible moral subjects. On one natural reading of Frankena's position, it is only *human* actions, desires and traits of character that can be evaluated from the moral point of view. Secondly, Frankena argues that we can morally evaluate one's actions, desires and dispositions, but we are not making a *moral* judgment when we, for instance, are making normative comments about one's looks, ethnic origin or one's linguistic accent. There are intuitively plausible constraints on what can be taken as a morally relevant feature of agents. Not every aspect of a person can be the proper object of moral assessment. Frankena's extended list of the morally relevant features can perhaps be reduced to *intentional actions* as the primary bearers of moral value.[15] On the other hand, one's intentions, desires, motives and character can be taken as morally relevant in a *derivative* sense only – their moral relevance is parasitic on the fact that they tend to affect and produce certain types of actions.[16] In a possible world where there is no connection between one's 'inner' state and one's outward behavior, the moral evaluation of one's intentions and desires would be quite out of place.

If actions are the primary bearers of moral value,[17] then one is taking a moral point of view if (minimally) one makes a normative judgment about one's actions or any other feature of the inner state or character that may affect or lead to actions. At the same time, a refusal to make such a judgment in relevant circumstances may (occasionally) indicate a non-moral stance assumed by an observer. But clearly not every instance of a judgment about human behavior warrants the name of a moral judgment. I may strongly disapprove of one's conduct without yet expressing a moral disapproval, as when I disapprove of the handling of the puck by a hockey player on the ice. In addition, Frankena argues, one must be willing to universalize one's judgment and be able to justify it in the right way.

The requirement of universalizability is a formal requirement on a moral judgment or principle, a constraint that deals both with the way it is (ideally) formulated and its intended scope. It is a requirement that excludes from the moral realm all evaluative formulations whose validity is essentially tied to the specifics of the place, time and circumstances in which a judgment is made. If one judges a certain action to be morally right or wrong, one is committed, given this requirement, to pass the same judgment on all relevantly similar actions as well. At the very least, it rules out mere numeric differences between the two actions or persons as irrelevant. More formally, as Mackie suggests, this implies that "a judgment containing a proper name or indexical term used not as a variable but as a constant will not yet be universalized," [13, p. 84.] and thus anyone making this kind of judgment is not making a moral judgment, but rather a non-moral one. In this sense, the requirement of universalizability amounts to the requirement of being sincerely willing to apply the same prescription or evaluation to oneself as well as to others, as long as there are no relevant qualitative distinctions between the cases.

This last condition is related to Wilson's 'Anonymity requirement' for moral judgments. She places the following formal constraint on the way moral rules are expressed:

> Prescriptive proposals, even if they arise from within particular cultural settings and reflect the concerns of creatures known to be partial to themselves, presuppose a detached perspective. There is an anonymity

> requirement on moral theorizing. […] The requirement implies that the endorsement and propagation of norms that differentially serve the interests of the particular reference class that endorses and propagates the norm qualifies as ideology, not moral theory proper. [20, p 23]

Both Wilson and Frankena require that moral rules are formulated and applied without bias, treating all similar cases and all similar subjects of moral evaluation in a similar way. The point is quite familiar, and it brings to mind the venerable image of a blind-folded justice, impartially weighting the evidence. Admittedly, there remains much vagueness as to what counts as relevant or irrelevant difference between cases or people, but, for our purposes, it is primarily important to notice what kinds of normative systems this requirement *excludes* from the set of rival moralities, relegating them to the class of personal policies, political or religious ideologies.[18]

The third condition from Frankena's account deals with the kind of *justification* that a proponent can provide for his normative judgment, granted that the judgment is about human actions and the proponent is willing to universalize it in the relevant sense. The justification condition is the central issue here. The first two requirements are still not restrictive enough to eliminate a number of *prima facie* non-moral rules and non-moral evaluative judgments, such as "it is despicable that Glenn does not cross his hands twice when entering a house." The judgment is about Glenn's actions, and the proper name used in this context is a variable that one might be willing to substitute with a universal quantifier. But, on this view, it is not a moral judgment unless one makes it for the right kind of *reasons*:

> It seems to me that what makes some normative judgments moral, some aesthetic, and some prudential is the fact that different points of view are taken in the three cases, and that the point of view taken is indicated by the kinds of reasons that are given. [3, p. 110.]

The right kinds of reasons, according to Frankena, are "facts about what the things judged do to the lives of sentient beings in terms of promoting or distributing non-moral good and evil." [3, p. 114] The non-moral good and evil can be variously specified, but it is traditionally explained in terms of pleasure and pain. The less tendentious formulation will employ the notions of well-being and ill-being, since the descriptive element contained in these two terms is marginal (as compared with 'pain' and 'pleasure' or even 'happiness' and 'unhappiness'), which in turn allows for greater flexibility in formulating a *substantive* theory in normative ethics.[19] Furthermore, Frankena specifies the possible recipient of the non-moral good as a 'sentient being', a category that encompasses humans, animals and (potentially) any other creatures capable of experiencing pain and pleasure.

There is, however, an important ambiguity in Frankena's characterization of the third condition, which will refer back to a similar ambiguity in Wilson's account as well. This ambiguity can be illustrated by the following two examples. In the first case, a witch-doctor makes an evaluative judgment "It is always wrong to spit in the direction of the Great Forest", and, when prompted, explains it by reference to the bad-tempered Evil Spirit, who lives in the forest, and who tends to get offended by these kind of actions. As a result, the Evil Spirit might destroy the crops by causing the drought, which will hurt the community. In the second case, a scientist claims that it is wrong to divert the river from its current course, and justifies his judgment by arguing that the lack of water will cause a drought, which will hurt the crops, and, ultimately, hurt the community. The relevant difference between the two examples is that, in the first case, the witch-doctor is clearly mistaken in his factual beliefs about the damaging effects of spitting, whereas in the second case, the scientist is probably correct about his predictions. Does the witch-doctor then make a genuine *moral* judgment?

More generally, we can ask whether the justification story for one's evaluative judgment in terms of welfare changes has to be also *accurate*, in addition to it simply containing references to

(sincere) *beliefs* about the alleged consequences of actions, in order for that judgment to qualify as a moral one.[20] The emphasis on the objective accuracy of the justificatory story would exclude from the set of all moral judgments those prescriptions and value judgments which are based on clearly mistaken factual premises; whereas the more liberal interpretation would recognize the prescriptive claims of the above-mentioned witch-doctor as moral pronouncements despite the falsity of the justification story. Before settling the issue, we should look at both alternatives more closely.

Ronald Milo, among others, embraces the latter, liberal interpretation of the third condition when he writes:

> If one holds that it is wrong to step on the lines of a pavement and if one's reason for thinking this is that this will cause great harm by bringing on a plague, then one may be said to accept this principle as a moral one. [15, p. 197.]

On Milo's view the thing that matters in deciding whether a person assumes a moral point of view or some other non-moral normative perspective is the subjective state of the person who makes a value judgment of some kind. In so far as he or she (sincerely) *believes* that stepping on the lines of the pavement will bring great harm, his proscription to do so assumes a moral status, and no further connection of those beliefs with reality is necessary.

This interpretation, however, cannot be accepted for the following reason. As we noticed earlier, the emphasis on the subjective side of justification of one's normative judgments threatens the descriptivist or essentialist account of morality to collapse back into neutralism. If one's beliefs about the facts how "the things judged" (e.g., actions) affect the welfare of sentient beings do not have to be true or even probable, but only, perhaps, sincere at most, then any normative judgment can be turned into a moral one by acquiring a necessary set of (false) beliefs about the consequences. It is conceivable that the normative claim "All members of the ethnic group X are evil, and thus should be exterminated" can be universalized in the required sense (even including the willingness to apply this prescription to the one who utters the claim if he turns out to be the member of this group), and justified by references to the alleged immense harm that the members of this group produce for the rest of the world. As long as the factual story is sincerely believed, on Milo's interpretation, it will be a peculiar *moral* judgment endorsing a genocidal policy. Yet, part of the motivation for embracing the descriptivist account of the nature of morality in the first place was precisely the promise that an account alternative to neutralism would exclude from the moral domain any system of action-guiding rules which incorporates this or similar prescriptions (e.g., 'Nazi morality'), and would further specify objective criteria for such exclusion. As the two recent authors keenly observe: "If you are able to honestly examine the moral arguments in favor of slavery and genocide, then you are likely to be either a psychopath or a philosopher." [5, p. 196] But, surely, they have meant a philosopher who adopted neutralism.

3. The Realism Constraint

If a descriptivist account is indeed to be a genuine alternative to moral neutralism, a more restrictive interpretation of the justification condition needs to be adopted. It should be formulated in a way that would exclude cases where the justification story behind a normative claim, which makes references to the alleged changes in welfare or an advantage-reducing effect of a certain prohibition, has no connection with reality. The modified interpretation, in other words, should respect the *Realism constraint*.[21] The Realism constraint, in the present context, brings a healthy dose of 'objectivity' into the otherwise purely subjective account of justification, which would refer only to one's beliefs, hopes and expectations about the outcome of observing a prohibition or following a certain rule.

I suggest, though, that it is neither necessary nor practical to demand that the justification of one's normative judgment in terms of (future) changes in the welfare of sentient beings as a result

of intentional actions must actually be true on *all* occasions when a judgment is made. Even if the above-mentioned scientist is accidentally mistaken in that particular case in his estimation of the adverse effect of the diversion of the course of the river, he is still making a genuine moral judgment, since it is presumably more likely than not that the substantial harm to the village will ensue as the result of these rash manipulations with nature. It seems sufficient to accept a probabilistic account of changes in welfare (or minimizing the disadvantages of the weaker subjects), using general experience and available scientific information as our primary guides in a specific situation. To be sure, all initial references to probability are inherently vague so that a number of borderline cases will always remain (e.g., *how* probable the alleged future harm needs to be in order for a prescription to qualify as a moral judgment, etc.). But adopting this interpretation will set certain objective constraints, and will at least allow us to say this much: Milo's example of "it is wrong to step on the lines of a pavement" will be a non-moral judgment in most contexts, whereas "it is wrong to let children play on thin ice" will be a *prima facie* moral claim, since, as long as we are talking about the actual world, the harm ensuing from playing on thin ice is much more likely to follow than the alleged plague that is believed to occur as a result of stepping on the lines of a pavement.

The proposed Realism constraint is meant to save a descriptivist account of morality of the type proposed by Wilson and Frankena from the collapse into moral neutralism. Admittedly, these are not the only ways of describing what a moral point of view amounts two, although they appear to be the most plausible ones. But it seems clear that any account of morality which harbors essentialist or semi-essentialist ambitions must built a similar constraint into its definition of morality.

A required bond between justification for a proposed normative rule or prohibition and the real world might be differently specified, with various degrees of strictness. On the one extreme, as we have seen, one might demand a perfect coincidence between the predicted effects of the rule-following (or rule-breaking) in a given case and the real set of consequences of one's actions. Yet, most would agree that this interpretation is too demanding, as it will exclude from the set of morally relevant rules or judgments the ones which warn of serious risks which (luckily) remain unfulfilled. "It is wrong to drive while under influence of alcohol" is a genuine moral evaluation for obvious reasons, even if on an occasion a drunk driver makes it home safely. Similarly, using Wilson's criterion, a proposed advantage-reducing imperative (e.g., a certain affirmative action policy) need not benefit the disadvantaged in all cases in order to count as a moral imperative. After all, a rule prohibiting using one's sexual charm while applying for a job may well leave an otherwise incompetent and, thus, a disadvantaged candidate, without the last hope of securing the employment.

Interpreting the Realism constraint in a too liberal way, on the other hand, would risk blurring the line between the moral and non-moral judgments, rules and prohibitions. If a judgment qualifies as a moral judgment as long as the justificatory story behind it is *sometimes* accurate, we open the door wide to a slippery slope regress, and, ultimately, to a reduction of normative sphere of morality to a purely descriptive discipline. Sure enough, it might well happen that someone, somewhere and for some reason could have gotten hurt when he saw a person stepping on the lines of a pavement, which would effectively bring Milo's rule into the moral domain.

We may also observe that the problem of application of the Realism constraint becomes less troubling as we move higher on the scale of generality of rules. "One should not exploit children's ignorance for personal gain or comfort" is a much more probable candidate for the moral domain than "One should not dismiss questions about sex when asked by one's children." The latter rule might well be a sound (or unsound) educational advice with no obvious moral implications, even though it can be plausibly derived from the former, more general prohibition.

What I hope to have done in this article is to have brought out certain deficiencies of the descriptivist accounts of morality and sketched one plausible way of overcoming them. The two versions considered are representative of the mainstream metaethical discussions in this area, and there good reasons to suppose that any alternative account of the moral point of view which as

much as recognizes that the effect of the rule-following on other sentient beings constitutes at least *a* relevant feature in determining its status as a moral rule, will face similar difficulties in spelling out the connection between the intended and actual consequences of a normative policy. The proposed Realism constraint is by no means a magic stick which would easily solve these problems – we have seen some inherent ambiguities in the constraint itself – but it is something that must be addressed by the philosophers who try to formulate the substantive theory of the nature of morality.

References

1. Benedict, Ruth. *Patterns of Culture*. (1935). Boston: Houghton Mifflin, 1989.
2. Bennett, Jonathan. "The Conscience of Huckleberry Finn." *Philosophy*, vol. 49, no. 188, 1974, pp. 123-34.
3. Frankena, William K. *Ethics*. Upper Saddle River, New Jersey: Prentice Hall, 1963.
4. Garner, Richard. *Beyond Morality*. Temple University Press, Philadelphia, 1994.
5. Haidt, J. and Bjorklund, F. "Social Intuitionists Answer Six Questions about Moral
6. Psychology." *Moral Psychology*, vol. 2, 2008, pp. 181-218.
7. Hare, Richard M. *Freedom and Reason*. Oxford: Oxford University Press, 1963.
8. Herman, Barbara. *The Practice of Moral Judgment*. Harvard University Press, 1996.
9. Hobbes, Thomas. *Leviathan* (1651). ed. Edwin Curley, Indianapolis: Hackett, 1984.
10. Joyce, Richard**.** *The Myth of Morality*. Cambridge: Cambridge University Press, 2002.
11. Kant, Immanuel. *Groundwork of the Metaphysics of Morals* (1785). Mary Gregor, trans. & ed.; intro. by C. M. Korsgaard. Cambridge: Cambridge Univ. Press, 2001.
12. Lucretius, Titus Carus. *On the Nature of Things* (1st cent. B.C.E.). Tr. by W. H. D. Rouse. Cambridge, Mass.: Harvard University Press, 1975.
13. Mackie, John L. *Ethics: Inventing Right and Wrong*. Penguin, 1977.
14. Mill, J. S. *Utilitarianism*. (1861). Roger Crisp ed., Oxford: Oxford University Press, 1998.
15. Milo, Ronald D. *Immorality*. Princeton: Princeton University Press, 1984.
16. Nielsen, Kai. *Why Be Moral?* Buffalo: Prometheus Books, 1989.
17. Sartre, Jean-Paul. "Existentialism." (1945) in *Existentialism and Human Emotions*, tr. Bernard Frechtman. New York: Philosophical Library, 1957: 9-40.
18. Turnbull, Colin. *The Mountain People*. Simon & Shuster, Touchstone edition, 1987.
19. Vetlesen, Arne Johan. *Perception, Empathy, and Judgment: An Inquiry into the Preconditions of Moral Performance*. The Pennsylvania State University Press, 1994.
20. Wilson, Catherine. *Moral Animals: Ideals and Constraints in Moral Theory*. Oxford: Clarendon Press, 2004.
21. Zavaliy, Andrei G. "Toward a Three-Dimensional Model of Moral Domain." *Southwest Philosophical Studies,* vol. 31, 2009, pp. 37-44.

Notes

1. The phrase 'demarcation problem' used in the present context first appears in Catherine Wilson's "Moral Animals" (2004), Chapter 1. But there is a well-known analogous problem of demarcating between science and pseudo-science (as opposed to making a distinction between good and bad science).
2. This problem of demarcation equally applies to moral *rules* and moral *issues* as well as to moral evaluative *judgments*. We may ask what makes the evaluative statement "This is a bad weapon" an instance of a *non-moral* judgment, as opposed to "He is a bad person", which is a *prima facie* moral judgment. We will still need to have some criterion to be able to make these distinctions. For brevity sake, I shall talk mostly about moral rules and principles but much of the same reasoning will apply to judgments as well.
3. We can mention the rules of prudence (e.g., "Always stay away from stray dogs"), the rules of etiquette (e.g., "Always hold your fork with the left hand") and the game rules (e.g., "Move your pawn in such and such a way") as examples of the *prima facie* non-moral rules.
4. Ethical Relativism, as a theory in normative ethics, presupposes the truth of moral neutralism, but is not identical with it.
5. It is possible to give both the *objective* and *subjective* formulations of neutralism. In the former case, it would be defined (as I did) in terms of empirically observed behavior of the agents, while in the latter case, the emphasis is given to the inner attitude of approval or disapproval that an agent acquires when presented with a principle of conduct or an occasion for a certain type of behavior (e.g., 'X is wrong' is defined as 'X rouses indignation in me'). As long as there are no references to content, form, goals or grounds of these attitudes, and 'whatever principle is (strongly) approved' is used as the criterion by which we distinguish between the moral and the non-moral, we still have a version of neutralism. Since it is natural to expect that the attitudes of approval or disapproval will be (generally) manifested in one's behavior, I shall mainly refer to conduct here.
6. My emphasis.
7. The anti-social and uncooperative patterns of behavior of the Ik culture in Uganda are documented by Colin M. Turnbull in "The Mountain People" (1987). The peculiar behavioral code of the Dobu culture of New Guinea was described by Ruth Benedict in her "Patterns of Culture" (1936/1989).
8. A short excerpt from Himmler's speech given in 1943 in front of SS audience may illustrate this: "We had the moral right vis-à-vis our people to annihilate this people which wanted to annihilate us. But we had no right to take a single fur, a single watch, a single mark, a single cigarette, or anything whatsoever…" (quoted in Vetlesen, 1994, p. 110). What is peculiar about this SS Weltanschauung is that the theft of a cigarette is morally wrong, but the collective annihilation of millions is part of one's moral duty.
9. For more detailed criticisms of neutralism see Milo, 1984, Chapters 6 & 7.
10. That is, criteria that are relatively stable and do not shift from individual to individual.
11. For reasons that will hopefully become clear, I violate the chronology and discuss Wilson's account before Frankena's.
12. Hobbes version of the origin of morality in the "Leviathan" (1651/1984) is probably the most widely known. But he is by no means the first philosopher to propose a similar account. Hobbes' main ideas can be traced back at least to the writings of Lucretius in the first century B.C.E. ("On the Nature of Things").
13. Admittedly, there are alternative interpretations of Kant's position, which tend to mitigate this requirement, especially with respect to imperfect duties (e.g., Barbara Herman, 1996). But nothing in the argument to follow depends on a particular interpretation of Kantian ethics.
14. The last, fourth, requirement, of course, is only conditionally necessary – it applies only to cases when the self-referential evaluative judgments are made.
15. The question what precisely makes *some* actions morally relevant events cannot be fully discussed here. I suggest that there are two interrelated features. First, an action is morally relevant if it is intentional, i.e., *free* in some acceptable sense of this word. Secondly, it is morally relevant if it is a type of event that may affect the wellbeing of others.
16. Kant famously argues for what appears to be the opposite view. For Kant, it is the will (which includes intention) that is good or bad in the primary sense, and actions have positive moral value only derivatively, in so far as they are the products of the good will. (Kant, I. 2001, p. 1) But Kant still recognizes the connection between good will and moral actions when he argues that the moral status of an action might be questionable *because* it might be the result of, say, selfish motives. His further claim that a good will would still "shine like a jewel" even if no actions were to follow seems counterintuitive.
17. 17. We do sometimes apply moral predicates to states of affairs, e.g., "this system is unfair.", but these statements can in principle be analyzed as referring to the actions of particular persons or groups of people.
18. This does not imply that any rule with the *less-than-universal* scope of application is a non-moral rule. Morality is essentially scalar, and certain rules may be 'more moral' than others. For an argument to this effect see Zavaliy, A., 2010. But it seems clear that the overly restrictive policies, such as personal ethical egoism, would not pass this test. As Kai Nielsen observes: "Personal ethical egoism isn't a malign, satanic or cynical morality because it isn't a morality at all. [P]ersonal egoism is a mere contempt for moral considerations

altogether and this is not a moral code at all. It is not even something that could be intelligibly and coherently proclaimed as a morality" (1989, pp. 158-9).

19. Hare objects that "to put this restriction upon the use of the term 'moral' is to write some kind of utilitarianism into its definition." (1963, p. 163). This seems to be an overstatement, since all that such content-based definition of morality implies is that the fact about human welfare must constitute at least a *relevant* reason for making a moral judgment. There is nothing in this definition of morality that specifies the weight one should assign to these reasons, or the extent of impartiality with which one should consider these facts. Both Kantian ethics and Utilitarianism are compatible with Frankena's understanding of morality defended here.
20. This question is not identical to the more familiar question of what should determine the *moral worth* of an action, its actual consequences or the intention of the agent.
21. Wilson does introduce a 'reality constraint' in Chapter 2 of "Moral Animals" (p. 49), but not as an element of the demarcation problem between moral and non-moral domains. The constraint helps Wilson to introduce the notion of a 'moral para-world' in order distinguish good or acceptable moral rules from poor or unacceptable rules (which still remain *moral* rules in the wider sense).