Moral Worth And Normative Ethics

Nomy Arpaly

WORK IN PROGRESS.

According to a number of theorists (Arpaly 2002, 2003; Arpaly and Schroeder 2014; Markovits 2010), a morally right action has moral worth if and only if it is performed for the right reasons, which are the reasons for which it is right, or the right-making features of the action. Actions with moral worth can also be said to be praiseworthy actions, and to be performed out of good will.¹

The idea central to these theories, of acting for right reasons through acting on right-making features of actions, deserves particular attention.

Right reasons for action do not need to be reasons that the agent believes to be the right reasons for action.² Huckleberry Finn, as I have argued, is praiseworthy for helping Jim if his reasons for helping Jim are the right-making features of helping Jim (that it treats Jim with respect, perhaps), however wrong he takes his action to be. Likewise, a person who does the right because she thinks it is right

¹ It should be added that what holds for right action also holds for supererogatory action, so really “morally desirable” might be more accurate than “right”. For ease of reference, though, I’ll talk about morally right actions.

² Arpaly (2003) defends this thesis at length.
might not be praiseworthy for her action at all, if she is acting for the wrong reasons. A person might think that the moral thing to do is to promote the interests of Aryan people over people of other races, for instance. If this person hires the better job candidate, not because she is better, but because she is of the “Aryan race,” she acts rightly, and thinks she acts rightly, but still for the wrong reasons, and so does not act in a way that is praiseworthy after all. She performs the right action, but not for its right-making features.

Another way to put this point, found in Michael Smith (1994), is to hold that acting out of a commitment to the right de dicto, as opposed to de re, is a kind of moral fetishism. On my view, being committed to the morally right whatever it is (when “maximize the beauty of ducks” is a possible candidate for the right moral theory in the agent’s mind) is not praiseworthy in the least. It is perhaps clearer (and, if different, better\(^3\)) better to say that a praiseworthy action stems from a commitment to the right conceptualized correctly. If utilitarianism has the right account of the features that make actions right then the agent performing a morally worthy action conceives of her action as increasing happiness, and is committed to increasing happiness so conceived; if Kantianism is correct she conceives of her action as respecting persons, and is committed to respecting

\(^3\) See Arpaly and Schroeder (2014, chapter 7).
persons so conceived; and so on. Whether she also conceives of her action as “the right action” is immaterial.

Very few people act on exactly the right reasons, however, (it is also worth noting that no one acts on exactly the wrong reasons or the reasons that make actions wrong – imagine a person who acts, for example, to minimize utility). Yet many people are praiseworthy for their right actions. So it should also be held that one can act in a praiseworthy manner without acting out of a commitment to the whole of the right. A person who does the right thing in order to increase (but not necessarily maximize) happiness generally, or even in order to increase the happiness of specific people or groups, should be held at least somewhat praiseworthy by a utilitarian, if not ideally praiseworthy. A person who does the right thing from a commitment to respect persons (so long as those persons are men, or Whites, or heterosexuals, etc.) is not worse than most people in history, including Kant himself. These intermittently respectful people should also be held praiseworthy to some degree by a Kantian.

What seems common to these two sorts of cases is that, in addition to there being complete or exhaustive moral reasons to perform actions, there are also partial, pro tanto moral reasons to perform actions. And a person who is not committed to
the complete moral reasons might be committed to at least some partial, pro tanto moral reasons, and act upon them while doing the right.

In this paper, my aim is not to defend the claim that for a right action to have moral worth it needs to be performed for the right reasons, which are the reasons for which it is right. I have defended this claim in a paper (“Moral Worth”) and a book (Unprincipled Virtue) and there is no need to repeat what I said. My aim in this paper is to explore some implications of the claim and see what other claims have to be accepted if my theory of moral worth is taken seriously. I find this interesting because if the connections between moral worth and moral reasons argued for by me and Markovits hold, if it is taken for granted that the morally worthy action is done for its right making features, then one can make inferences from claims about moral worth to claims about what the right making features of actions are, and so to claims about which normative ethical theories might be correct.

Markovits (2010) has already effectively applied the idea to utilitarianism in arguing that, if utilitarianism were correct then people who keep promises for non-utilitarian reasons would act with no moral worth. One way to put it is to say that she would act with no more moral worth than Kant’s prudent grocer. Kant’s prudent grocer, recall, acts honestly because he reasons that acting
honestly will bring him money. He is a paradigm of a person whose right action has no moral worth. But not everyone who keeps promises for non-utilitarian reasons is equivalent to him in the moral worth of his or her actions. Yet, if utilitarianism were correct, it would follow that these people, just like the grocer, act for morally irrelevant reasons and not for the right-making features of their actions. And so, they would act without moral worth, in a way that was not in the least praiseworthy (however instrumentally convenient it might be to praise them). This, Markovits argues, is a reason to reject utilitarianism. It is of course open to the utilitarian to reject the intuition that philosophically unsophisticated people act with moral worth for non-utilitarian reasons rather often, just like it is open to the utilitarian to reject the intuition that it’s wrong to hang an innocent man to prevent a race riot. There is always a choice, but Markovits, I think, has a convincing case when she shows the costliness of one such choice. I do not wish to argue that in every case in which a normative claim (such as utilitarianism) clashes with an intuition about moral worth we should follow the latter, but sometimes the price of conflicting with an intuition about moral worth can be too high for a normative theory.

Prima facie, if Markovits and I are right about moral worth, utilitarianism is not the only theory that has problematic implications for moral worth. Take neo-Aristotelianism, or at least a simple version thereof. I am referring to the view
that the right action in any circumstances is the action that would be performed by the virtuous person under these circumstances, a virtuous person is a person who has the virtues and these, in turn, are character traits that promote flourishing. Imagine asking a person “why did you keep your promise?” or “why did you help the person bleeding by the side of the road?” and getting the answer “because this is what a person would do whose character is conducive to flourishing.” There is something a little strange about this answer: it is too much about the life of the agent and too little about the life of the person to whom the promise was made or who was found bleeding by the side of the road. It seems as if it is the flourishing of the person bleeding by the side of the road that should be on the agent’s mind, not the agent’s own flourishing, much less that of a hypothetical virtuous agent. I also wish to point out, in the spirit of Markovits’s attack on utilitarianism, out that many praiseworthy agents seem to act from motives that have nothing to do with flourishing.5

The neo-Aristotelian might reply that the virtuous person does not act in order to gain the good life or even in order to do what a virtuous person would do. She

4 See, e.g., Hursthouse (1998)
5 Note that this is a criticism of Neo-Aristotelianism and not of Aristotle’s own view: while Aristotle would agree that the right thing to do and what the virtuous agent would do are the same thing, he never clearly says that the right thing to do is the right thing to do because the virtuous agent would do it. He never holds that defending your city in war is a fine thing because it would be done by a person whose character is conducive to flourishing, where the “because” indicates the right-making feature of the act. In fact, he never tells us at all why defending your city in war is a fine thing, so it is impossible to criticize his view of the right-making features of defending your city, unless we take it to be that defending your city is right simply because it is defending your city.
might reply that the virtuous person helps the person bleeding by the side of the road because the person needs help. But completely divorcing the motives of the virtuous person from the right-making features of her action has strange consequences. As per the neo-Aristotelian, what makes an action right is the fact that it follows from a character trait that is a component of the good life. If the virtuous person’s motive for helping has nothing to do with these right-making facts then it seems to be the case that either the virtuous person’s actions have no moral worth – after all, it’s only an accident that she did what the flourishing-prone agent would do, as nothing was further from her mind – or the right-making feature of helping the person bleeding by the side of the road does consist of the fact that the person needs help rather than in facts about the flourishing-prone person. It could still be true that one performs a right action if and only if one does what the flourishing-prone agent would do, but this equation would not give us the essence of what makes an action right.

In the next section I will discuss some implications that truths about moral worth would have on normative ethics if the Arpaly-Markovits view is true. In the following section I will argue at some length for one particular such implication: that there is more than one kind of moral reason, or, in other words, more than one right-making feature for actions. This claim, I will argue, follows from facts about the moral worth of altruistic action, a topic on which Kantianism is wrong.
General Implications

What are the right reasons for a moral action? I think that, when asking that question, we need to take it as a starting point that there are actions that have at least some moral worth. Kant doubted whether any morally worthy action had ever been performed, but I see no need to share his brand of pessimism. One thing to remember about agents who have acted in a morally worthy manner is that most of them are not ethicists. I will not refer to these agents as ordinary people, as some of them are quite extraordinary, but they are not all philosophically sophisticated. It is unlikely that when they act well they follow, with any precision, the categorical imperative or the principle of utility (or Peter Railton’s specific consequentialism,\textsuperscript{6} or...). That alone is not enough to rule out either the categorical imperative or the principle of utility as moral truths because it is possible that most praiseworthy moral actions are committed for \textit{pro tanto} moral reasons rather than complete moral reasons. However, the reasons for which the philosophically unsophisticated praiseworthy agents act should be complete moral reasons, or be \textit{pro tanto} moral reasons, or at the very least be trivially derivable from such. They must have something to do with the true normative theory: to be the whole of it, or part of it, or at the very least a trivially

\textsuperscript{6} See, e.g., Railton (1988).
derivable consequence of it. Commonsense morality is not necessarily the right moral theory, but some commonsense moral reasons – the ones acted upon by praiseworthy agents – are at least pro tanto moral reasons.

What moral reasons do philosophically unsophisticated praiseworthy agents, whether ordinary agents or philosophically unsophisticated saints and heroes, act from? I will not try to provide an exhaustive list of reasons that philosophically unsophisticated praiseworthy agents seem to act for, but it is easy to think of some. “I promised” is one such reason: quite often people explain why they kept a promise simply by appeal to the fact that they have promised. The fact that someone is in need is cited just as often as a simple reason for acts of benevolence. In all likelihood these are reasons that genuinely motivate many praiseworthy acts of promise-keeping and of benevolence, respectively, and thus a true normative theory must account for their status as complete moral reasons or pro tanto moral reasons for action, or at least trivially either derivable instances of such reasons. For the moment, take for granted that I promised and someone is in need are good candidates for moral reasons – at least pro tanto moral reasons for action.

The Arpaly-Markovits thesis, again in conjunction with the fact that there exist praiseworthy agents who are philosophically very unsophisticated, has another
implication for the search for the right reasons for moral action. These reasons, and therefore the right normative theory, should not be such that the rationale behind common, uncontroversial right actions require an amount of reflection or cognitive sophistication that only a few possess. For lack of better words, it is an advantage for a moral theory not to be too lofty. Now, one has to be careful not to underestimate the amount of complexity of which people are capable without being able to articulate this complexity. People are capable of learning complicated dances without being able to provide purely verbal instructions as to how to perform them, and they are capable of learning complicated languages without articulating their grammars (but this special ability to act for grammatical reasons, an ability learned in such a special way, is probably not representative of our abilities to act for complex reasons in general). However, there is something suspicious about a moral theory that justifies uncontroversially right or good courses of action in a manner hard to explain to anyone but the brightest students. For example, explaining to all but the brightest college students what “contradiction in conception” means is remarkably hard, and so I hold in suspicion the view that people who are not the brightest students routinely act to avoid such a contradiction. One must note quickly that this criterion does not rule out Kantian reasons in general. After all, considerations such as that you would not like it if everyone did X, or that you would object if someone did X to you, considerations that are at least Kantian in
spirit, are understandable even to preschool children. When I say “uncontroversial” moral action I am allowing for the fact that some moral actions, such as voting for the right candidate in politics, do require reflection due to their very nature. One must not expect all morally praiseworthy agents to be able to articulate their reasons for keeping promises or helping a bleeding person by the side of the road but it speaks well for one’s theory if one’s description of these reasons does not require a university-level course to understand.

One last general implication of this view of moral worth is the following. If it is true then saying, “I thought this action would be the right one but I was wrong” is not in itself giving a good excuse, even when wholly sincere and correct: a person who performs an action because she think it right can still display monstrous indifference to the wrong-making features of the action. However, if you perform the wrong action because of innocently believing that it had some right-making features, where your idea of what would make an action right is the right idea, it seems as if you should have an excuse. Thus, “she thought it would save this person’s life but she was wrong” would, if true, excuse an agent from blame (so long as this being wrong amounts to an innocent mistake rather than self-deception or culpable ignorance). Thus, a moral theory of the form “the right thing to do is the action that has property X” has to hold that a person

7 See Arpaly and Schroeder (2014), chapters 8, 9.
would not be blameworthy if the agent were to have performed a morally unfortunate action out of an innocent belief that the action would have property X. Thus, if you want to say that the right action is an action that an impartial observer would recommend, you are committed to the excuse that, “he thought that his action would be recommended by an impartial observer.” As a matter of fact, this is a pretty bad excuse, and so it would seem that even if it is true that an action is right if and only if it would be recommended by an impartial observer, the fact that it would be recommended by an impartial observer cannot be its right-making feature.

To take a more substantial example, consider the fact that an action respects autonomy. This is a much more plausible candidate for the right-making feature of an action, but subjecting it to the innocent mistake test shows that our knowledge of the nature of autonomy is still too vague for this to be the right-making feature of an action itself. Usually, “she thought her action respected autonomy but she was wrong” would not be seen as a simple excuse for assisting suicide by the person who thinks that assisting suicide fails to respect autonomy, nor would it be seen as a simple excuse to support a law against assisted suicide by the person who thinks that doing so fails to respect autonomy. For each side, the agent on the other side would appear not as respecting autonomy while being factually mistaken but rather as disrespecting autonomy. She has acted for
a reason that she conceived of as respecting autonomy while it was something else. Unlike saying, “it would save his life,” saying, “it would respect his autonomy” is still too vague a statement for mistakes about autonomy to count as a regrettable innocent mistake that does not reflect poorly on the moral worth of the actor. The innocent mistake test presents an ideal of clarity that a moral theory should aspire to meet, and that most current ones do not.

Altruism

In this section I would like to defend the view that actions can have moral worth when they are the result of concern for the wellbeing of others. I will also argue that, if we assume the Arpaly-Markovits view of moral worth, it follows that there are at least two kinds of moral reasons, altruistic reasons being one such kind. In other words, there are at least two different right-making features that different actions can have, one of which is the fact that an action protects well being, or, if you wish, prevents or alleviates illbeing. The distinction between alleviation of ill-being and increasing wellbeing is blurry, but I will talk mostly about ill-being as the clearest duties we have that involve other people’s well being are duties to protect wellbeing – to prevent or alleviate illbeing (rather than duties to increase the wellbeing of people who are already doing well). I will not offer a position on the question of whether what is going on here is that we have
a greater duty to help a person the worst off she is, though this is what I suspect. There is no precise word in English for concern to protect wellbeing. There are words like ‘compassion’, ‘kindness’ and ‘sympathy’, but they only refer to different emotional manifestations of such concern. In the mean time I will refer to it, following Lawrence Blum (2009) as “altruism.”

A person acts from altruism, in my sense, to the extent that she acts in order to protect wellbeing in general or protect the wellbeing of a particular person because it is wellbeing. In other words, she acts out of an intrinsic concern that people not suffer illbeing or from an intrinsic concern that a particular person or particular people not suffer illbeing, which is a realization of a concern that people not suffer illbeing: when acting out of altruism you care about the protecting the well being of a fellow human being (the various altruistic emotions such as compassion and sympathy happen when you pay attention in some way to the fact that someone is a fellow human being and similar to you in relevant ways, and note that the person is not doing well). When I say that one concern “realizes” another I am referring to cases in which you care about something because it instantiates something else which you care about intrinsically. For example, my concern that Tim not have the flu realizes my concern that Tim be healthy. When I say that a concern is intrinsic I mean the one cares about something non-instrumentally and also not as a realization of another concern. . I
refer to the protection of human wellbeing not in order to deny, necessarily, the view that one should substitute “sentient being” for “human being” but because when I argue with Kantianism, which I will do presently, I would like to argue that Kantianism fails to capture some of our intuitions even if we restrict our discussion to protecting human wellbeing.

If actions are morally worthy because of being performed for their right-making features then there are some actions for which altruism is not a moral worth-granting motive. Promise keeping is the right thing to do because of features other than the fact – when it is a fact – that it protects wellbeing. However, it is natural to suspect that some actions, such as helping the person bleeding by the side of the road, are right because they reduce illbeing, and thus altruism is a morally worthy motive for them. Pre-theoretically, many people would regard a person who helps another person out of pure altruism as a person who is praiseworthy for her action. If Jennifer helps John, who has serious difficulties writing his dissertation, despite the fact that he is not her student or her close friend, because she wants to alleviate John’s illbeing, we regard her as having performed a morally worthy action. She is not analogous to Kant’s prudent grocer: helping in order to alleviate illbeing is not analogous to helping for a material reward or in order to have a nice story for the novel one is writing.
According to Kant and Kantians, a person acting out of altruism simpliciter performs an action that does not have moral worth. When I say altruism “simpliciter” I refer to a concern with the alleviation of ill-being that does not include the extra consideration known as the avoidance of a contradiction in one’s will. Kant famously grounds the obligation to help people in the idea that if one wants to be helped, also willing never to help creates a contradiction in one’s will. Christine Korsgaard (1996,60) argues that the morally worthy helper, according to Kant, has a motive to help that is deeper than that of the sympathetic helper because he has a “further stretch of motivating thought.” That further stretch “concerns the sort of world that this would be if no one helped, or better still if no one perceived the need for help as a reason to help (…..) Such a world would be unacceptable because we regard our own needs as reasons why we should be helped.”

I doubt many philosophically unsophisticated agents ever think of what might happen “if no one perceived the need for help as a reason for help”, but let us grant that Korsgaard’s consideration has some guise in which it can occur to a philosophically unsophisticated agent. I take it, however, that pre-theoretically many people would not require this consideration to be in the background of an altruistic right action for it to have moral worth. It seems that we would be inclined to praise the helping actions of a being for whom wanting to be helped,
or, if you wish, the avoidance of contradictions in the will, is not part of her reason for action, for the simple reason that there would be no contradiction in her will if she were not to help. The literature on Kant sometimes mentions a person who does not want help. Imagine again Jennifer, who helps John, who is having trouble finishing his dissertation, despite the fact that she has no special obligation to him. Imagine that Jennifer, a proud Canadian, responds to all offers of help with the words “no, thanks, I’m fine.” She helps John, but she herself, would never accept such an offer of help: she does not wish anyone to spend resources on her that can be spent on some good cause. Jennifer’s will would not contain a contradiction if she were to decide not to help John, as she would not like to be helped herself. Yet, if Jennifer helps John for no ulterior motives, because a person is suffering ill being, her actions appear praiseworthy. Hence, some praiseworthy helping actions are performed without regard to contradictions in the will or the fact that one wants other people to help one.

Kantians could argue several things. They could argue that Jennifer’s attitude for which she refuses help is morally objectionable – that she ought to have more respect for herself as help-worthy and regard her needs as reasons for people to help her. If that is true, though, it could still be the case that her help of John is morally worthy. She would be a person who lacks a certain virtue that can be called self-respect but who does, at least, perform some morally worthy helping
actions. Kantians could also argue that I have not taken seriously enough the
need to universalize Jennifer’s maxim, and we must remember that Jennifer, if my
example is to work at all, is not only a person who would not accept help if she
were in John’s shoes but a person who would not accept help at all. A person
who would not accept help at all, the Kantian might say, is necessarily irrational,
because no human fails to have ends for which it would at some point be useful
to her to receive help, and whoever wills the end wills the means, if she is
rational. It is hard to adjudicate as to how to generalize Jennifer’s maxim, and
people also have ends to which not being helped is instrumental (e.g., the end of
becoming independent). However, the Kantian may insist that any human
being’s ends are such that never receiving help is, all things considered, irrational
for them to commit to. Nonetheless, I am not sure that Jennifer cannot be
irrational in her thought that she does not need help (or in her not wanting it)
and morally praiseworthy for helping. A person can be praiseworthy and foolish.
Leaving that aside, however, I would like the reader to think of a being who does
not need help.

Imagine a non-human Jennifer who does not need help, perhaps because she has
magical powers that help her out whenever she would, if she were human, need
help. Imagine that the being in question helps people and does not do so to

---

8 I thank James Dreier for this point
benefit herself or for any other reason but to alleviate illbeing. The fact that the being herself does not need or want any help means that there would be no contradiction in her will if she were to decide never to help anyone, and so she is in no way acting in order to avoid a contradiction in her will. However, it is easy to imagine the angelic super-being’s actions as praiseworthy. In fact, several people to whom I spoke claimed that a being who does not need help is in one sense more praiseworthy for helping than a creature who does need help. Why? It is not clear. Perhaps because her empathy with creatures so unlike her is impressive. Or perhaps – and this might be the most promising explanation – because she does more than is fair for her to do.

Some would say that even if the non-human creature, with her magical powers, were not to act on the universal law formula she might be acting from the formula of humanity. I will turn to this thought later in the paper.

It is worth noticing that acting out of altruism, in some of its guises, such as compassion and kindness, can look surprisingly similar to acting from obedience to the categorical imperative. After all, for a concern that people not suffer illbeing to produce a concern for the alleviation of a particular person’s ill-being there is typically a need for the agent in question to be struck by the humanity of the sufferer. Ideally, the altruistic person cares as much for the alleviation of
another’s illbeing as much as she cares about the alleviation of her own suffering, and so she shares with the person who follows the categorical imperative a certain bracketing of the distinction between self and other. She might well (sometimes) be advised to imagine herself in someone else’s shoes before she acts, simply because she needs to have a good idea of what the other person feels, wants, and needs. All of these things, however, do not amount to aiming at universalizability in the Kantian sense.

Let us go back to the controversial Kantian thesis that helping actions motivated by altruism *simpliciter* have no moral worth. Why would one accept it?

The main reason to think that actions motivated by altruism simpliciter have no moral worth is the same reason Kant (and most of us) think that the right actions performed by Kant’s prudent grocer do not have moral worth: because it appears that altruism only accidentally leads to right action. There are various reasons to suspect this is the case, and I will started with those I consider less good. It is possible to use the term “altruism simpliciter” very broadly, or just differently, and to attribute it to agents who are taken by the large eyes of a baby or a fluffy rabbit and feel an atavistic warmth towards the large-eyed creature that might result in an urge to alleviate the illbeing of the creature in question. I do not wish to argue that these people’s actions, to the extent that they are motivated by
cuteness alone, have moral worth, as the cuteness of the object of compassion has nothing to do with the right-making features of helping her. These cuteness-susceptible people are to true altruism what people susceptible to authoritative voices are to a real sense of duty. The truly altruistic person is not affected by a lack of cuteness in the person before her, in the same way that a person truly respectful of autonomy is not affected by the fact that the person before her is ditzy and nervous. For the purpose of this paper we shall ignore responses to cuteness, whether interpreted as atavistic responses planted in us by evolution so as to promote good parenting or interpreted more cynically by Kant when he says, “We love everything over which we have a decisive superiority, so we can toy with it, while it has a pleasant cheerfulness about it: little dogs, birds, grandchildren”.

Babies and rabbits aside, why would one think that altruism only accidentally leads to right action? One answer that is sometimes attributed to Kant is that the altruistic person acts from an indirectly selfish motive: she enjoys making people happier, and helps them for that reason. This need not be true. An altruistic agent might care about protecting wellbeing for its own sake. In that case, any pleasure or relief from pain that she receives through helping people is just a side effect of that intrinsic concern. If a person does not care about the victory of the red sox than she does not get any pleasure out of red sox victories. Similarly, if the agent
did not have an intrinsic concern for the alleviation of ill-being she would not have felt pleasure at the alleviation of other people’s ill-being. Like the Aristotelian virtuous agent, she might get pleasure out of her action but she does not act for pleasure (assuming that she gets any pleasure at all: contrary to Kant, one can care about the well being of others even when one is too depressed to feel pleasure). She need not act selfishly.

A similar charge to that of hedonism, though a more complex one, is the charge of narcissism. The Kantian might say that the altruistic agent helps “because she feels like it”. She has a desire to help, or she cares about helping, and she values or desires doing whatever satisfies her desires or pertains to her cares – whether it happened to be protecting well being, promoting the sales of British Shorthair cats, or whatever else catches her fancy. But again, that need not be the case – even the person who intrinsically desires the victory of the Red Sox need not be such a narcissist about her desire. A person might greatly hate herself, putting no value on the satisfaction of her desires qua her desires, but still desire or care about the victory of the Red Sox, taking such a victory to be a good thing despite, rather than because of, the fact that it’s her team. Similarly, if Robert desires Martha’s wellbeing and helps her as a result, it need not be the case that Robert is motivated, when he helps, by the thought “Robert wants Martha’s wellbeing”. Robert, who might, as a self-hater, not care very much about Robert and what he
wants, need only be motivated by the thought “Martha needs help”. Acting on a desire does not mean acting on a desire to fulfill your desires or a principle that says “fulfill your desires!”

A better reason to suspect that altruism only accidentally leads to right action is because it seems not to lead to it reliably. Not all reliable motives for moral action are by definition non-accidental – the prudent grocer’s actions would not have moral worth even if a sort of invisible hand guaranteed that self-interest always led to right action – but all non-accidental motives to right action are, presumably, reliable, and altruistic inclination, says the Kantian, is not.

However, not all types of unreliability imply accidentality. The fact that an altruistically inclined person might lose her concern for protecting wellbeing does not show that her motives to right action are accidental, that her reasons for action are not the reasons that make it right. After all, sadly enough, dutiful people can lose their sense of duty. That is true even if by “dutiful people” we do not simply mean those who feel a sense of duty to do what they think is right but rather we mean those who feel a sense of duty when and only when a duty actually exists, even if we interpret “a sense of duty” as “susceptibility to golden-rule type reasoning”. A sense of duty can come and go. The fact that a sense of

\[9\] See Pettit and Smith (1990)
duty is something that one could lose does not entail that acting from a sense of
duty cannot be acting for the right reasons. The same is true for altruism.

But other ways of being accidental are quite relevant to our discussion. A
motive’s connection to the right action appears accidental if the motive can lead to
wrong action. This is one reason I hold a desire to do what is right, conceived of
as what is right, not to be a moral-worth granting motive: such a desire can
motivate one to do bad things, as one might think that the wrong things are right
and so have a misguided conscience. Famously, some have argued that altruism
is unreliable in this particular way - it leads to wrong actions sometime\(^\text{10}\). One
obvious example would be paternalistic lies. This kind of argument should be
particularly bothersome to me, as it seems to imply that my own semi-Kantian
view of moral worth implies that altruistic actions don’t have any.

Consider again Jennifer, who helps John with his dissertation in order to alleviate
illbeing. Imagine that after getting to know Jennifer a little better you discover
that she has at some point lied to her roommate. Jennifer, it turns out, faced the
situation described by Thomas Hill in “Autonomy and Benevolent Lies.”\(^\text{11}\) Her
roommate had barely recovered from an extremely painful love affair with a

\(^{10}\) Though I rely a lot on Barbara Herman (1992) as the provider of possibly the best arguments
on behalf of the Kantian view of moral worth, I do not intend to use her example about helping
the art thief struggling with his load, as there is a complication in the fact that the altruist in the
example appears to be, as one undergraduate put it to me, “an idiot”.

person with whom she was obviously incompatible. One night she asked Jennifer whether she thought her ex-boyfriend would be open to getting back together, and Jennifer said “no,” despite knowing that the ex-boyfriend had, in fact, wondered aloud about the rightness of the breakup and expressed a desire to get back together. Suppose Jennifer said “no” to her roommate for the same reason for which she helps John with his dissertation: she wants to alleviate illbeing. Jennifer’s action, I agree with Hill, is wrong.

Knowing that about Jennifer and her roommate, would one reevaluate her action vis-à-vis John and his dissertation and regard it as morally worthless – akin to the actions of the prudent grocer? In the world outside of philosophy, few people would do that. Some would take the dissertation case as evidence that even in telling the lie, Jennifer was well-intentioned. While it is not clear what “well-intentioned” means in such a case, it is worth noticing that no one calls Kant’s prudent grocer well-intentioned. What is going on here?

The following theory can explain our intuitions. There is more than one type of moral reason – more then one right-making feature that different actions can have- and it is possible to act in a way that reveals a commitment to only part of morality. There are at least two types of moral reasons – one that has to do with wellbeing and one that has to do with something else. Perhaps there are more
than two. At any rate, altruism fails to deliver right actions every time not
because it resembles, in terms of moral relevance, the motives of the prudent
grocer but because sometimes moral reasons are present that override reasons of
altruism. Considerations that have to do with wellbeing present a moral reason
to help that is sometimes overridden by other moral reasons, reasons to do
something else.

Some would suspect that the reason many people do not count Jennifer’s lie to
her roommate as a reason not to see moral worth in her act of helping John with
his dissertation because her lie seems out of character, but it need not be. It
might be, but it does not matter: the point of the story is that Jennifer, when she
helps John, acts on a motive that in at least one case led her astray. But let us
assume that the lie is in character. Jennifer is altruistic through and through,
which she exhibits through many right actions but also the occasional overly
paternalistic action, always in the kind of circumstances present in the roommate
case, i.e., circumstances in which a seemingly minor paternalistic action can save
a person a lot of illbeing. Character evaluations and moral worth are tightly
related; when we regard a person as having good character we do not just regard
her as someone who can be relied on to perform good actions, as a self-interested
person can be thus reliable if she is convinced that acting well will bring her
good things or if she lives in a friendly, “invisible hand” world. A person with
good character is a person who responds well to moral reasons. Jennifer’s paternalism appears to be a character flaw, but a character flaw in a person who also has some respects in which she is good. Again, we would say “Jennifer is well-intentioned but there are some things she just doesn’t get.” In other words, Jennifer seems to be a person who is indifferent, or at least fails to respond sufficiently, to one type of moral reason, despite responding fully to another. Jennifer’s case indicates the existence of more than one kind of moral reasons.

What if, instead of the gently paternalistic Jennifer, we were to imagine a Stark Raving Utilitarian? The SRU would kill one to save five, frame an innocent man to prevent a race riot, and maybe do some other things that the stereotypical utilitarian boogieman would do, all in the name of reducing ill-being. He is sincere in his utilitarian motives, and would be ready to sacrifice himself as much as he is willing to sacrifice others. It has been pointed out to me that we would never say about the SRU, after counting his sins, “at least he is kind” – something that we might say about Jennifer. Yet the SRU seems different only in degree from Jennifer, whom I regarded as a person who responds to one type of reasons and not another. If Jack is also acting for one type of reasons but ignoring another type, why don’t we say that he is at least kind, even if he is not just?
There are two reasons for that. One reason consists of a fact about kindness. The SRU is, indeed, unkind. The person for whom the word “kindness” is reserved in English tends not only to be concerned with alleviating illbeing but also has other characteristics. She is, at least, deeply averse to being the cause of illbeing, especially extreme illbeing. An aversion that deep functions as a kind of soft side-constraint against causing illbeing above a certain threshold. It might be possible to convince someone who is stereotypically kind to cause extreme illbeing but it cannot be easy. Note that in my previous example, Jennifer does not cause any illbeing with her lie, and so she does not demonstrate a lack of kindness. The SRU, on the other hand, does cause illbeing through his Stark Raving Utilitarianism. There is no word in natural English (“altruism” being a term of art) to describe both Jennifer and the SRU because, as I have said, there is no word in English to describe concern for the alleviation of ill-being (The Sanskrit word karunā does, apparently cover both agents. It is often translated as “compassion”, but karunā is attributed to the Buddha when he calculatingly kills a group of people who would otherwise commit a sin that would doom them to a miserable afterlife, a case in which using the word “compassion” appears to be a stretch.12)

12 Thanks to Nic Bommarito for these facts about Sanskrit.
Another reason for rejecting “but at least’s he is kind” has to do with the implicature of saying good things about bad people and with the essence of blameworthiness. Blameworthiness (as understood in this paper) is negative moral worth. If a right action has moral worth to the extent that it is performed for moral reasons, a wrong action has negative moral worth if the agent, in performing the action, exhibits indifference to moral reasons, or, in some cases, performs the action for the very reasons that make it wrong. Indifference to moral reasons come in degrees. One can take moral reasons into account but not enough, for instance. The SRU exhibits monstrous indifference to at least one type of moral reasons, and the monstrosity of his indifference is great enough that it sounds bad to say anything good about him at all. Above a certain threshold of badness, saying anything good about a person or a thing – especially if prefaced by “at least” – is problematic. It is a fact about our linguistic practices that most people would not say, except in very cynical moods, that the Spanish civil war was bloody but at least it was fascinating. We would certainly not say such a thing to a Basque whose relatives were killed by Franco’s troops. However, the Spanish civil war was, and is, fascinating, and the SRU, while not kind, is altruistic. He is different, in terms of his moral profile, from a person who commits the same crimes for morally neutral motives like money or power, though his victims will hardly be consoled.
When the SRU does something good in order to alleviate illbeing, does his action have moral worth? Well, sometimes. If the right-making feature of the action in question is the fact that it alleviates illbeing – say, giving most of his money to a charity that fights famine, or risking himself to fight a dictator who tortures his people- his action has moral worth. If, on the other hand, he keeps a promise purely because it alleviates illbeing, then his action has no moral worth, because the right-making features of keeping promises are not related to well-being.\(^{13}\)

Some would object that the person who kills one to save five cannot possibly be doing something that she has even a *pro tanto* reason to do, simply because she is doing something wrong, something forbidden by a side constraint, and even something that a virtuous person would never consider doing. On my view, he does something that he has a *pro tanto* reason to do (save five people) even as he does something that he has a *pro tanto* and an overriding reason not to do (kill one person). In order to accept that there is a side constraint that forbids the action and that it would not occur to a virtuous person to do it one does not need to argue, like McDowell (1979), that there are never *pro tanto* reasons to perform wrong actions. I myself am convinced by Seidman (2005) that there can be such a thing as a *pro tanto* reason to do something which is in fact wrong, even if there cannot be overriding reasons to do wrong things. Seidman provides us with an

---

\(^{13}\) Though by keeping a promise the SRU might also prevent some ill-being, and of course that would be morally worthy in itself. Sometimes one has both promise-based and altruism-based reasons to keep promises.
example of a businessman who is looking desperately for ways to avoid bankruptcy. Since he is virtuous, it does not occur to him to take immoral courses of action for that purpose. He does not consider having his creditors killed, for example, nor does he consider cooking the books. Yet, when we praise the businessman for allowing himself to lose his business rather than do the immoral things that someone else might have done we acknowledge that he would have had a reason to do these things. Otherwise, there would have been no sacrifice involved at all, and it would be hard to see where the admiration came from that we feel for those who give up something valuable for the sake of doing the right thing. The businessman does sacrifice something when he allows himself to go bankrupt, and he is justified in being sad when he loses his business. Similarly, I think, the person who virtuously refrains from killing one person to save five people sacrifices something valuable – in this case, morally valuable, as are the lives of five people – that cannot be regarded as trivial, as is suggested by the fact that it makes sense for him – in fact it is morally virtuous for him - to be profoundly sad after the five die, sad both because of the deaths themselves and because of the fact that he had to refrain from helping the five.

If Jennifer’s action of helping John has moral worth then the protection of wellbeing, or at least the alleviation of ill-being, is a moral value, something we have pro tanto reasons to pursue as moral actors. If there are in fact at least two
types of moral reasons for action – at least two right-making features of actions – than there are at least two independent moral values. A true moral theory will have to imply that there are at least two things that make actions right, that there are at least two moral values that well being is one of them. From a Kantian point of view, some people might argue that the moral duty to help others is but an outgrowth of the formula of humanity of the categorical imperative: the duty to further the ends of other persons. But the duty to further the ends of other persons, at least as normally conceived by Kantians, is incompatible with paternalism, and Jennifer’s motive, when not constrained by other motives, leads her to paternalism. Though ordinary English makes it sound strange, Kantianism regards paternalism such as the sort expressed in benevolent lies as a way of treating persons as means, or at the very least as a way of disrespecting the human ability to set ends by making it harder for the person paternalized to set her own ends. Thus, Jennifer is not, on common interpretations of Kant, acting from the formula of humanity. More obviously, the SRU is not acting from the formula of humanity when he gives all of his money to a charity that prevents famine, as his motive leads him, unconstrained, to do things that the formula of humanity precludes. The value of alleviating ill-being seems to be separate from the value of treating people as ends in the specifically Kantian sense.

Works Cited


Seidman, J. “Two Sides of Silencing”. *The Philosophical Quarterly* Vol. 55, No. 218 (Jan., 2005), pp. 68-77