"Contemporary Kantian Ethics"

Andrews Reath

Introduction: Some Main Themes in Kant's Ethics

Kant's project in ethics is to defend the conception of morality that he takes to be embedded in ordinary thought. The principal aims of his foundational works in ethics – the *Groundwork of the Metaphysics of Morals* and the *Critique of Practical Reason* – are to state the fundamental principle of morality, which he terms the "Categorical Imperative", and then to give an account of its unconditional authority – why we should give moral requirements priority over non-moral reasons – by grounding it in the nature of free rational agency. Roughly the principle of morality gets its authority from the fact that it is by acting from this principle that we exercise our free agency. In these works Kant develops a distinctive account of the content of moral requirement (which is filled out in his later work, *The Metaphysics of Morals*). According to one version of the Categorical Imperative, we determine what sorts of actions are permissible or required in various situations by asking whether a principle of action is rationally willed as universal law for agents with autonomy. A second version of the Categorical Imperative derives the content of morality from the principle that we are to respect "humanity", or "rational nature", as an "end in itself" and never merely as a means. "Humanity" is the capacity for autonomous rational choice, and it includes the capacity to act from one's own judgment of what one has reason to do, to set ends for oneself, and to guide one's actions by values one finds it reasonable to accept. To hold that this capacity is an end in itself is to claim that it has an absolute value – a value that Kant terms "dignity" – that sets limits on the

proper treatment of and is the basis of positive duties toward any agent with capacity. Since, in Kant's view, all normal individuals possess this capacity equally, it grounds the fundamental moral equality of all individuals, in virtue of which they are owed moral concern. The principle of treating humanity as an end in itself and never merely as a means thus bases the content of morality in respect for persons as rational agents with autonomy. It leads to an ideal of moral community in which relations between persons are based on mutual respect for autonomy, an ideal that Kant terms a "realm of ends". Since Kant thinks that rational agents necessarily value the capacity for rational choice and the related capacities for self-governance, this approach to the content of morality also suggests a justification of its authority.

Kant's overall moral theory also suggests a meta-ethical stance. While he thought that moral principles are objective requirements based in reason, he rejects the moral realism of his rationalist predecessors which holds that there are metaphysical facts about right and wrong that are part of the nature of things and whose truth is independent of the operation of our practical reason. He also rejects empiricist accounts that base moral judgments in features of our psychology, e.g. holding that properties of virtue and vice are the tendency of certain actions to elicit feelings of approval or disapproval when we consider them from a general and impartial point of view. Rather, Kant held that the fundamental principle of morality is based in the nature of rational volition, and that particular moral requirements are based in principles that we autonomously impose on ourselves through reason. This aspect of his moral thought is seen in the idea that we determine the content of morality at some level of generality by determining whether a principle of action can be willed as universal law for agents with autonomy, or whether it

shows proper respect for humanity as an end in itself – that is, through rational procedures based on commitments that we have as rational agents. Some commentators have termed this feature of Kant's moral theory a form of "moral constructivism": moral requirements do not reflect an order of moral truths that are part of the nature of things, but rather are "constructed" by an idealized rational procedure. The truth or objectivity of a moral principle is explained by the fact that it is justified through this kind of reasoning.

Thus Kant suggests distinctive answers to several different questions in moral theory: what are the basic standards of right and wrong, and what are the fundamental values that drive moral thought? What sorts of claims are moral claims – what are they about and what determines their truth or falsity – and how do we establish them? What explains the special authority of moral requirements? Contemporary Kantian theories, while they may modify or reject many of Kant's specific views, follow Kant's lead in developing answers to some or all of these questions. This article will give an overview of Kantian approaches to the content of morality, Kantian constructivism, and Kantian accounts of the authority of morality. But first we begin with a note on Kantian moral psychology.

Kantian Moral Psychology

Humeans believe that motivation is desire-based – that reasons for action and motivation are ultimately traced back to desires and fundamental preferences that arise in an agent independently of any practical reasoning. If I desire to take a trip to Tahiti and I need to set aside money to afford it, then my desire to take this trip can give me a reason to begin the regime of savings. Furthermore, the realization that saving money is a means

to my end can redirect my desire for the end toward the means, thus producing the motivation to begin saving. Because Humeans believe that all motivation is desire-based, they explain moral motivation in terms of some natural desire or psychological mechanism, such as sympathy or natural concern for others, or a tendency to identify with the well-being of others.

Contemporary Kantians reject the Humean view of reasons and motivation because they believe that moral principles are requirements of reason that apply to agents independently of desire. They are committed to holding that human beings can be moved to act by reason alone. Kantians hold that it is part of rational agency that one can be motivated to act by one's application of rational principles and one's judgments about what one has reason to do, without the intervention of any desire or further source of motivation. In the above example, the fact that I need to begin saving money in order to afford my trip is a reason to begin saving, and the judgment that I ought to begin saving money now by itself can motivate me to do so. Likewise the judgment that I ought to take steps now to ensure my well-being later in life can motivate me to do so, without any further felt desire. (Note that the claim is that one *can* be motivated by one's judgment of what one has reason to do – that is not to say that one always *will* be motivated by that judgment.) Since the reasons in these two cases ultimately stem from some desire (e.g., some future desire), the full significance of the Kantian view of motivation comes to light in moral cases. Here Kantians hold that moral requirements apply to us simply as rational beings independently of our desires, and that the judgment that we ought to perform (or refrain from) some action can motivate us to do so, without the stimulus of any further desire. So for example, judging that I ought to refrain from taking unfair advantage of a

competitor or that I ought to help someone in need can motivate me to do so. The Kantian view here is that the application of principles of reason (or the judgment about reasons) produces the motivation to comply with the principle and does not simply redirect or elicit a prior motivational state that exists independently of any reasoning.

Kantian Approaches to the Content of Morality

The component of Kant's moral theory that has had the most influence on contemporary normative ethics is the principle of treating persons as ends in themselves and never merely as a means, and the related ideal of relations between persons based on mutual respect for autonomy. Many theorists have thought that this principle, suitably developed, can ground at least significant portions of the standards of right and wrong.

The Kantian principle may at first seem to say only that we should not "use people" for our own purposes (or "use them without their consent"). But since we use the actions, decisions, and services of others all the time in morally innocuous ways, often without explicit their consent, it must mean more than this. We get more mileage out of the principle by noting that Kantian autonomy is a capacity for self-determination and self-governance that includes the capacity to form one's own judgments about good and bad reasons. Intuitively, we respect the autonomy of individuals, so understood, when we allow their use of these capacities to set limits on how we may treat them. This thought suggests an ideal of justifiability to others, and that is how the principle of respecting persons as ends in themselves is now widely understood at the most general level: to treat persons as ends and never merely as a means – that is, to respect persons as autonomous self-governing agents – is to act from principles that others can freely endorse (as agents

with autonomy) and that justify one's actions to them. Or it is to act from reasons that others can reasonably be expected to accept. "Justifiability to others" implies that the justification of an action is to be addressed to those affected by it as rational agents, in light of their fundamental interests. I cannot expect others to endorse or to accept my reasons and principles unless they acknowledge other people's equal moral standing and give adequate weight to their fundamental interests, including their interest in exercising autonomy and self-governance. Since actions that do not meet this standard are off limits, this ideal gives individuals a kind of hypothetical veto power over how others may treat them. In this way, the ideal of respect for the autonomy of individuals, when specified through the idea of what can be justified to others as agents with autonomy, leads to strict principles of conduct that recognize persons as moral equals.

This ideal of justifiability to others needs to be understood in a strongly non-consequentialist fashion. The specific principles to which it leads set limits on how one may promote desirable outcomes or overall good. They may require some action even when an alternative produces a better overall outcome. Furthermore, the reasons identified by these principles are not simply weighed against competing reasons (such as those based in the desirability of some outcome), but rather can silence them or undercut their force. For example, consider a situation in which some action furthers a desirable outcome, but fails to satisfy the general criterion of justifiability. Perhaps one can advance one's career through an act of deception that will undermine someone else's prospects; or perhaps violating someone's legal rights, or torturing them, may further the security of one's community. In these cases, one cannot expect the person on the receiving end to accept these particular ways of furthering one's ends. (As an agent with

autonomy, that person has no reason to endorse these ways of pursing one's ends, since they infringe the person's capacity for self-determination and self-governance.) Normally the fact that an action may promote some good (one's career, national security, etc.) is a reason in its favor. But in these circumstances, that fact has no force as a reason. In other words, one does not just weigh reasons that favor the action based on the desirability of the outcome against reasons that oppose it stemming from its failing the criterion of justifiability. The fact that the action violates an individual's autonomy undercuts and excludes the force of any reasons based in the desirability of the outcome.

Finally, since the resulting principles do not aim at promoting or maximizing some value (such as individual autonomy), they do not underwrite certain forms of reasoning, for example, that some action should be chosen because it leads to fewer overall infringements of autonomy in individuals or because it produces more opportunities for autonomy across individuals. Rather, these principles are required by the ideal of justifiability to others or respect for persons as agents with autonomy. This gives us a way to understand how actions can be right or wrong in themselves. Of course, action on such principles standardly aim at some outcome (protecting an individual from harm or aggression, providing aid that will preserve a person's capacity to exercise her agency, etc.). But they are understood as forms of concern that are owed to persons as such.

The general requirement of justifiability to others as equal autonomous agents readily translates into familiar more specific moral principles. It leads to requirements to avoid or refrain from gratuitous injury, coercion, deception and fraud, manipulation, exploitation and profiting from the weaker position of others, and so on. The rationale is

not simply the generally harmful effects of such actions on individual well-being. Rather, these kinds of actions infringe on individual's capacities for self-governance and self-determination, and autonomous individuals can reasonably object to such treatment. For similar reasons, it leads to a requirement to avoid paternalistic interference. It leads to requirements to refrain from free-riding and similar forms of unfairness, and to requirements of fidelity, keeping one's word, and not violating trust. Individuals who violate these requirements make an exception for themselves (by failing to do their share in cooperative schemes) or disappoint expectations that they have invited others to form. Because such actions fail to respect others as moral equals in various ways, they are not based on principles that equal autonomous agents can be expected to endorse. Finally, the ideal of justifiability to others grounds positive duties such as beneficence and mutual aid, gratitude, loyalty, special obligations between loved ones and friends, and so on, because such principles are among the social and material conditions needed to support the exercise of rational agency in socially interdependent beings.

In the political sphere, the ideal of justifiability leads to a liberal theory of justice, such as that developed by John Rawls (Rawls 1971, 1999). Rawls's theory guarantees all citizens a set of equal basic liberties (such as liberty of conscience, freedom of expression and association, rights of political participation, and so on) and substantively equal opportunity to compete for positions of social and economic advantage. His "difference principle" limits social and economic inequalities to the condition that they benefit those who are worst off. Rawls understands the basic liberties and opportunities as social conditions needed for citizens to develop and exercise various rational and moral powers (a capacity for a sense of justice and a capacity to develop and pursue a conception of

their own good) and to participate fully in social life. The principles of justice taken together establish a framework in which individuals have both the constitutional guarantee and adequate material resources to exercise these capacities and to pursue meaningful conceptions of the good. When the principles of justice are satisfied, the social order can be justified to all citizens, even to those who are worst off.

These moral and political principles are shared by many normative theories. What distinguishes a Kantian approach is the underlying rationale: they are requirements not because they promote some set of values or good outcomes, but because they express respect for persons as equal autonomous agents and are a condition of relations between persons based on mutual respect.

While many people find this approach to the content of morality compelling, disagreements remain. Consequentialists, of course, insist that standards of conduct be tied to the promotion of certain values or good consequences. Other theorists worry that Kantians overvalue individual autonomy, or that they adopt an overly rationalist picture of human beings. In response to the latter set of worries, it is important that for Kantians autonomy is not, fundamentally, the ability to act on one's preferences whatever they may be, but a capacity for rational self-government that includes the capacity to form one's own judgment about reasons. Its theoretical role is to ground the equal moral standing of persons and to set standards of justification for action and social and political arrangements. Further, like any capacity, the capacity for rational self-government needs to be developed (through socialization, moral education, or interacting with others), it is often realized partially or imperfectly, and it can be diminished by adverse circumstances

(psychological, social, or material). Thus, Kantians need not deny the social interdependence of human agents.

Another question is that if rational capacity is the basis of moral standing, can Kantians accord moral standing to children, in whom the capacities remain undeveloped, or to the mentally disabled who will never develop or have lost the capacity? And what about the status of animals and nature? Children are easily included in the Kantian moral universe by noting childhood is a stage in the life of a person who normally develops to autonomy. Children need not be treated as adults (e.g., paternalistic intervention is warranted), but the proper treatment of children should keep in view the rational capacities that they are in the process of developing. Most Kantian theorists hold that the moral standing of mentally disabled or incapacitated human beings comes from their membership in a species in whom rational capacity is the norm. Again, the standards of proper treatment differ from those for fully competent agents, but they will require giving adequate weight to the interests of such individuals and respect for any level of self-government of which they are capable. Regarding animals and nature, to hold that autonomy confers special moral standing on persons is not to deny that there are other forms of value or that other kinds of entities that can make claims on us. Thus Kantians can allow that certain ways of treating non-rational creatures and entities are morally deficient.

Kantian Constructivism

Many contemporary theorists influenced by Kant have suggested deliberative procedures for arriving at moral principles that assign a role to reasonable choice or

agreement – what principles individuals with a concern for justifiability would choose or agree to. Rawls argues for his principles of justice by showing that they are the rational choice in the "original position" – a construct designed to represent an ideally fair choice between free and equal moral persons tasked with selecting principles of justice to govern their social order. T.M. Scanlon's contractualism explains judgments about right and wrong in terms of principles that could not reasonably be rejected by people motivated to find general principles of conduct acceptable to others with similar motivation. (Scanlon 1999) This suggests a method of identifying moral principles that asks whether individuals could reasonably object to being treated in certain ways (i.e., would have reason to reject principles that permitted such treatment). Thomas E. Hill, Jr. has suggested, as a modification of Kant's notion of a "realm of ends", a framework that assumes idealized Kantian legislators overridingly committed to respecting human dignity who are concerned to give all agents opportunity to exercise their rational capacities and live as autonomous agents, and who seek principles that all can endorse. The idea is to guide reflection about moral principles by asking what universal principles for regulating conduct such Kantian legislators would choose for themselves. (Hill, 2000: 33-57) These theories are all forms of moral constructivism – theories that derive the content of morality from an idealized process of rational deliberation.

The term "constructivism" was introduced into moral theory by Rawls, who defines it as the view that moral (and political) principles may be represented as the outcome of a "procedure of construction" that incorporates the relevant standards of practical reason – that is, a process of deliberation aimed at reasonable agreement along the lines described above. (Rawls, 1996: 89-90) Constructivism is an approach to the

justification of moral principles, holding that the rational acceptability of a set of principles is established by showing that they are what individuals would choose for themselves or agree upon through idealized rational deliberation. Constructivists hold further that the "procedure of construction" is the criterion of what is right: that what makes a moral principle correct is that it would result from this idealized process of deliberation. Constructivism contains an approach to meta-ethical questions about the nature of moral claims, the epistemology of moral principles, and their objectivity and truth conditions that is accepted by many contemporary theorists influenced by Kant. Constructivism need not be Kantian. What marks Kantian forms of constructivism is that the process of deliberation is structured not just by self-interested rationality (e.g., the "agents of construction" are not simply concerned to secure their own interests), but is constrained by the aim of reasonable agreement on principles that give due weight to the interests of all, or by the aim of justifiability to each individual.

The distinctive force of constructivism is best illustrated by contrasting it to the account of moral truth and objectivity given by certain forms of moral realism. Moral realists such as the rational intuitionists accept the existence of an order of moral facts or moral properties that is independent of our methods of thinking about them, analogous to a realm of mathematical facts, that can be known or grasped through rational reflection. The objectivity of moral claims comes from this mind-independent order of moral facts: moral judgments about right and wrong are true when they accurately reflect these facts. Constructivists, by contrast, do not appeal to any such mind-independent order of moral facts to ground moral objectivity. A moral judgment is correct not because it accurately reflects the independent moral facts, but because it is arrived at through correct reasoning

– i.e., through deliberation that satisfies the constraints that come from practical reason and the aim of reasonable agreement. Thus constructivists can hold that correct moral principles and moral facts are constituted by practical reasoning, in the sense that correct principles are those that would result from this idealized process of reasoning. Facts about right and wrong are a function of the application of these moral principles.

One feature brought out by this last point is that constructivism denies a sharp distinction between our epistemological access to moral principles and their truth conditions. We justify or come to have knowledge of a set of moral principles by seeing that they result from the idealized deliberative process, and of course we can only do that by deliberating as conscientiously as we can. According to constructivism, that process of reasoning also specifies their truth conditions: what makes a principle true or correct is that it follows from correct reasoning.

Constructivism is often described as a form of "moral anti-realism", but that label is misleading insofar as it associates constructivism with non-cognitivism. First, constructivism is not a form of subjectivism or relativism, since the standard of justification is not actual agreement on principles, but ideal agreement. Like cognitivists, Kantian constructivists hold that claims about moral principles can be correct or incorrect, and that they can be the objects of belief and knowledge (and not just the expression of pro-attitudes). Here it is important to realize that constructivism offers a distinctive account of objectivity that is an alternative to various forms of moral realism.

Kantian approaches to the authority of morality

Many people think that moral reasons have special authority in that they apply to us independently of our desires and take priority over competing reasons in cases of conflict. The ideal of respect for persons outlined above has great plausibility on its face, but is it a source of reasons that are necessary and inescapable? Some theorists argue that the evident appeal of some substantive value suffices to explain the authority of morality. Scanlon, for example, argues that the authority of morality can be explained through the value of 'mutual recognition' – the value of living with others on mutually justifiable terms. (Scanlon 1998: 162) Kantian approaches to the authority of morality go further by grounding it in inescapable requirements of rationality. It is an open question whether such arguments succeed, but they take us into deep questions about practical reason and agency.

One such argument has been made by Thomas Nagel. (Nagel 1978). Nagel tries to establish a rational requirement of "altruism" – the principle that the interests of others give us direct reasons for action – by basing it in a metaphysical conception of oneself as one person among others equally real. Rational individuals ascribe value to the satisfaction of their own needs. But recognizing the reality of others requires that what one asserts about oneself can be meaningfully asserted of others, thus that the satisfaction of others' needs has the same value that one ascribes to one's own. An individual who does not acknowledge this requirement of altruism in effect denies the obvious metaphysical truth of the equal reality of others.

More recently, Christine Korsgaard has attempted to ground the authority of morality in the conditions of rational agency. One component of her view is that the formal principles of practical reason – including that of acting from principles that can

hold as universal law – are "constitutive principles of rational agency." (Korsgaard 2009) That is to say that it is by following these principles that we constitute ourselves as authors of our actions, and that these principles are tacitly involved in all exercises of rational agency. If so, they are not coherently rejected by any rational agent. To see how these principles are constitutive of agency, consider that action is determining oneself to be the cause of some end. Thus rational action requires a self over and above one's various motives that chooses which motive to act on, and action is the work of the self as a whole, rather than of some force within the self. Korsgaard's argument is that these conditions are achieved by following the formal principles of practical reason. (One makes oneself the *cause* of some end by taking effective means to one's ends, and one makes *oneself* the cause of some end by acting from principles that can be willed as universal laws. We focus here on the second argument.)

First, a rational action is guided by some principle and for an action to be the work of the self as a whole, it must be based on a principle with which one identifies. So self-determination requires identification with some principle of choice. Second, the principle of choice with which one identifies must be a universal principle applying to a range of similar cases. To see why, imagine an agent that identifies with a "particularistic principle" with no implications beyond the case at hand. This agent would wholly identify with the present motive of action. But in this case, no distinction can be drawn between the self and the various motives within the self, and there is in effect no active self at work. Such a "choice" fails to satisfy a basic condition of action – it fails to constitute an active unified self – and would not count as volition. (Korsgaard, 2009:

4.4.3) Thus, self-determination involves giving oneself some universal law, and if so, the principle of acting from some universal law is a constitutive principle of action.

Since this is a weak principle that sets almost no limits on one's concrete principles of action, further argument is needed to get the basic principle of morality. A second component of Korsgaard's argument is that rational action involves valuing oneself as a person. Korsgaard holds that reasons for action are based in a "practical identity", or "description under which you value yourself" (Korsgaard 1996: 3.3.1) – that is, a self-conception based on such things as social roles, ties to others, personal ends and projects, and so on, through which one finds certain activities to be worthwhile. Many of our practical identities are contingent, and their hold on us is a matter of our continuing to endorse them. But since we need reasons in order to act and reasons are based in specific practical identities, the need for some practical identities is not contingent. That fact about rational action is the basis of a necessary identity – one's identity as a human being, "a reflective animal who needs reasons to live and act" (Korsgaard 1996: 3.4.7) Our human identity gives us higher order reasons to endorse and to take seriously some practical identities, and in acting on these higher order reasons one values oneself as a human being. Thus, the hold on us of our particular practical identities comes in part from our human identity (our need for some practical identities), and valuing oneself as a human being is a condition of having reasons and finding anything to be worth doing. In endorsing a specific practical identity, one also endorses the reasons that come from one's human identity and values oneself as a human being. (Korsgaard 1996: 3.4.7-9) What follows is that the conditions of agency commit us to valuing ourselves as human

beings, and valuing oneself in that fashion commits one to acknowledging the value of all human agents.

To combine these two arguments: you constitute yourself as an agent by giving yourself some universal law. Rational agency also commits you to valuing yourself as a human being in a way that acknowledges the moral standing of others. But if valuing humanity is implicit in all rational choice, then the universal laws by which you constitute yourself as an agent are implicitly laws for rational agents as such, or laws that all rational agents can endorse. In this way, Korsgaard argues that the basic principle of morality gives the basic form of action. This argument grounds the authority of morality in autonomy in the sense that the fundamental principle of morality expresses the conditions of rational agency.

KANT; REALISM; COGNITIVISM WITHOUT REALISM; INTUITIONISM; CONTRACTUALISM; RECOGNITION, RESPECT.

REFERENCES

Korsgaard, Christine M. (1996) The Sources of Normativity, Cambridge/New York: Cambridge University Press. (Referred to by section number.)

----- (2009) Self-Constitution: Agency, Identity, and Integrity, Oxford/New York: Oxford University Press. (Referred to by section number.)

Hill, Thomas E., Jr. (2000) Respect, Pluralism, and Justice, Oxford/New York: Oxford University Press.

Nagel, Thomas. (1978) The Possiblity of Altruism, Princeton: Princeton University Press.

Rawls, John. (1971, Revised Edition 1999) A Theory of Justice, Cambridge: Harvard University Press.

----- (1996) Political Liberalism, New York: Columbia University Press.

Scanlon, T. M. (1998) What We Owe to Each Other, Cambridge, Harvard University Press.

FURTHER READING:

Herman, Barbara (2007) Moral Literacy, Cambridge: Harvard University Press. (Explores Kantian approaches to several issues in moral theory.)

O'Neill, Onora (1996) Towards Justice and Virtue, Cambridge/New York: Cambridge University Press. (Develops a Kantian constructivist account of universal principles of justice and virtue)

Velleman, David (2006) Self to Self, Cambridge: Cambridge University Press. (Essays by a contemporary philosopher that explore Kantian themes.)