

Kant's Conception of Personal Autonomy¹

The Kantian view on the contemporary understanding of autonomy is ambivalent. This ambivalence arises because, in the contemporary literature, a strong distinction is often made between personal autonomy and moral autonomy, with Kant associated *only* with the latter. Joseph Raz, for example, writes: "Personal autonomy, which is a particular ideal of individual well-being, should not be confused with the only very indirectly related notion of moral autonomy. The latter originates with the Kantian idea that morality consists of self-enacted principles".² Personal autonomy, in the contemporary sense, involves (roughly) governing yourself in the pursuit of your own conception of the good, a conception that is usually stated in morally neutral terms. Personal autonomy does not, however, "(merely) [involve] unconstrained ... choice", since personal autonomy can be absent even when unconstrained choice is present.³ For example, a woman who consents to a cosmetic medical procedure as the result of brainwashing may not be exercising her personal autonomy, even though her choice is unconstrained.⁴ A focus on personal autonomy therefore leads to important questions about when it counts as *really you* doing the governing of yourself in terms of *your own* conception of *the good*. In contrast, moral autonomy involves (roughly) legislating the moral law for yourself. This leads to important, but different, questions about moral obligations and how morality limits the pursuit of the good.

Viewed in this way personal autonomy seems at best marginal and at worst a positive hindrance to moral autonomy, since personal autonomy may conflict with moral autonomy. Given that Kant is closely aligned with moral autonomy, can there be any legitimate role for personal autonomy (in its contemporary sense) within Kant's ethical framework? One strategy for finding such a role is to argue, as Jeremy Waldron does, that Kant's conception of happiness does the work of a theory of personal 'autonomy'.⁵ But the problem with this strategy, Waldron claims, is that Kant regards the pursuit of happiness as heteronomy and not autonomy, and heteronomy is incompatible with autonomy.⁶ So it seems that this strategy won't work. An alternative strategy is defended by Robert Taylor who argues that Kant's account of the obligatory ends of self-perfection and the happiness of others can do the work of a theory of personal autonomy.⁷ But this strategy relegates to the status of heteronomy what is normally

considered central to personal autonomy, namely the pursuit of our own merely permissible personal ends as part of our own conception of the good. So this strategy also seems to fail. Is there an alternative? I shall argue that there is by pursuing the following four goals in this paper.

The first goal will be to reject the claim that personal autonomy, properly understood, has no legitimate role in Kant's ethical framework by drawing out of Kant's work a distinctive unitary conception of autonomous willing. This will be a *unitary* conception of autonomy since it will cover both *moral* and *personal* autonomy as part of the one general account. This will help to achieve the second goal, which is to show how moral and personal autonomy can be understood as complementary to, rather than in tension with, one another. The third goal will be to help to address a gap in the literature by better situating Kant's conception of personal autonomy *within* (rather than simply *against*) the contemporary personal autonomy literature. Raz, as we have already seen, thinks that Kant has no conception of personal autonomy. Similarly, Onora O'Neill understands Kant to be defending a conception of "principled autonomy" which is incompatible with contemporary conceptions of "individual [or personal] autonomy".⁸ I will seek to counter such assertions by showing that Kant should be understood as endorsing a distinctive conception of personal autonomy that has many similarities, but also some important differences, with contemporary accounts of personal autonomy. This will lead to the fourth goal, which is to emphasise the important role that, according to Kant's conception, socialisation can play in both fostering and hindering personal autonomy.

This paper proceeds as follows. Firstly, we shall draw out of Kant's work a conception of personal autonomy. Secondly, we shall examine the role that socialisation plays in that conception. We will then be in a position to, thirdly, situate Kant's conception within the contemporary taxonomy of theories of personal autonomy before, fourthly, providing further examples to better illustrate Kant's theory.

1. Autonomy of the Will and Autonomous Willing

There are various things that we might call autonomous, such as, wills, willings (or choices), principles, states, and processes. Our focus here is on *personal* autonomy, that is, the

autonomy of persons, which is also Kant's main focus. Indeed, Kant links having autonomy of the will with *being* a person.⁹ But Kant writes of autonomy both as a *property of the will* and as a practical *principle*.¹⁰ How can it be both? The principle of autonomy is the principle that the will is "subject *only to laws given by itself but still universal*", and this principle, Kant claims, is also the supreme principle of morality.¹¹ Autonomy is a property of a will, then, just in case it is a will that is both subject to the principle of autonomy and capable of regarding itself as the lawgiver of that principle. Autonomy of the will requires the *capacity* both to regulate yourself in accordance with reason and to identify with those rational regulations. But a rational being must have these capacities in order to *be* a rational being. Autonomy of the will is therefore something which Kant thinks that *all* rational beings must practically assume that they and other rational beings have.¹² As Onora O'Neill puts it, "autonomy [of the will] is not the special achievement of the most independent, but a property of any reasoning being".¹³ Autonomy is opposed to heteronomy. Heteronomy of the will is the denial of the claim, strongly, that the wills of rational beings have the property of autonomy or, weakly, that human beings are rational beings in this sense. Kant rejects heteronomy of the will in both its strong and weak versions and argues that this view is the source of spurious principles of morality.¹⁴

But *autonomy of the will* is not the same thing as *willing autonomously*. This is because a will could have the *property of autonomy*, since it is subject to and able to regard itself as the lawgiver of the principle of autonomy, without *actually governing itself* in accordance with that principle, that is, without *willing autonomously*. This is why Kant contrasts "autonomy of the will" [*Autonomie des Willens*] with "heteronomy of choice" [*Heteronomie der Willkür*].¹⁵ Implicit in this contrast is Kant's crucial distinction, which we will return to later, between the will in its *legislative* function (*Wille*) and the will in its *executive* function (*Willkür*).¹⁶ As such, we need to clearly differentiate between autonomy as a property of *wills*, which is a matter of self-legislation or legislative function, from autonomy or heteronomy as a property of individual *willings* or *choices* [*Autonomie der Willkür*],¹⁷ which is a matter of self-government or executive function. Whereas autonomy of the will is a property of the will of *every reasoning being*, autonomous willing is an *ideal* achieved only by those rational beings who actually succeed in governing themselves in accordance with the principle of autonomy. If we are to draw a

conception of personal autonomy out of Kant's work, then we shall need to focus on autonomous willing or choosing, and not on autonomy of the will.¹⁸ This is because personal autonomy is generally understood to be an ideal achieved through making choices in a certain way, rather than as a property of the will of every rational being.

2. Autonomy and the Realm of Ends

To develop a conception of autonomous willing it will be helpful to briefly turn to Kant's account of autonomy in *Groundwork II*. There Kant argues that the principle of autonomy leads directly to a "very fruitful concept dependent upon it, namely that of a realm of ends" [*Reiche der Zwecke*].¹⁹ The realm of ends is a normative ideal made up "both of rational beings as ends in themselves and of the [personal] ends of his own that each may set himself".²⁰ The realm of ends is therefore made up of two types of ends: all existent persons, as ends in themselves, and the lawful self-given personal ends of those persons.²¹

In this realm all "rational beings stand under the law that each of them is to treat himself and all others *never merely as means* but always *at the same time as ends in themselves*", and each can "regard himself [or herself] as lawgiving" in relation to that law.²² Kant also identifies autonomy of the will as the "ground of the dignity of human nature and of every rational nature".²³ This dignity has an unconditional worth, since there is no condition under which it has no worth, and an incomparable worth, since the worth of persons, as ends in themselves, is not scalar and cannot be compared with the worth of things or personal ends.²⁴ Because Kant takes autonomy of the will to ground the dignity of persons, I will henceforth speak in terms of the more intuitive ideal of respecting the dignity of persons or treating persons as ends in themselves, rather than in terms of the more opaque ideal of self-given but still universal laws.²⁵

How do personal ends become part of the realm of ends? Each person must be treated as an end in itself prior to, and as a limitation on, whatever personal ends members choose to adopt. When a member actually wills or adopts a personal end which is compatible with that law, then her willing of that personal end introduces that new end into the realm of ends. In that case her will has *normative authority*, or is *legislative*, for all other members of the realm

of ends, that is, for all other rational beings. Personal ends which are introduced into the realm of ends therefore have an objective worth or worth for all rational agents.²⁶ But that worth is not unconditional (since its worth is conditional on its being introduced into the realm by a member) and is not incomparable (since its worth can be compared with other personal ends but not with the dignity of persons). This is another way of saying that the realm of ends formulation can ground an imperfect duty to make the permissible ends of others our own.

This means that others should regard *me* achieving *my lawful ends* as objectively valuable *because* they are lawful ends that I have set for myself. For example, suppose I adopt the end of painting a landscape this weekend. Assume that my willing this end introduces it into the realm of ends, since willing this end is compatible with treating all persons with dignity. You now have reason to help me to achieve my end by, for example, buying me a canvas to paint on. But my willing this end for me doesn't (at least normally) give you a reason to paint a landscape yourself. What might give you a reason to paint a landscape yourself is the fact that doing so would give you pleasure or contribute to your self-perfection. What my willing this end does is to make it the case that you have reason to help me to achieve my end. What it does not do is make it the case that you have reason to achieve that same end yourself. Further, you would not even be helping me to achieve my end if you were to paint a landscape yourself. What I will is not that landscapes be painted, but that I paint them, and that you recognise me willing this end as having normative authority for us. If you only respect what I will for some other reason, and not because you take my willing it to be authoritative for us, then you are not respecting my dignity as an autonomous agent. Of course, you also have reasons to help others to achieve their ends and reasons to achieve your own ends. That is why helping me to achieve my ends is objectively valuable but, unlike the dignity of all rational beings which you are always required to respect, not unconditionally or absolutely valuable. However, this only tells us how willing an end introduces it into the realm of ends. But how does one not merely will an end, but will an end autonomously?

3. How to will an end autonomously

Kant argues that "if the will seeks the law that is to determine it *anywhere else* than in the fitness of its maxims for its own giving of universal law – consequently if, in going beyond itself, it [the will] seeks this law in a property of any of its objects – *heteronomy* always results. The will in that case does not give itself the law; instead the object, by means of its relation to the will, gives the law to it".²⁷ This means, says Kant, that when we will an end autonomously we must "abstract from all objects [of the will] to this extent: that they have no *influence* at all on the will".²⁸ Otherwise a "foreign impulse would give the law to" our will and this is "*always only heteronomy* of the will".²⁹ This seems to imply that a person wills autonomously only when she *completely* detaches herself from the influence of her own desires and emotions, as well as from all social and even causal influences, and somehow wills an end without these things *at all influencing* what she wills. Does personal autonomy on Kant's account really require such complete transcendence? If so, then this might seem to be either an impossible or an unappealing conception of personal autonomy. Indeed, this would seem to vindicate the common view that Kant's account of autonomy is limited *only* to moral autonomy and that he considers what now passes for so-called personal autonomy (including the setting of non-moral ends) to be always nothing more than mere heteronomy. However, I shall argue that Kant's account of autonomy need not commit him to this view. To see why this is so, we shall need to look at Kant's account of willing in more depth.

To will an end is to take up an effective (in that it actually moves you to action) first-personal commitment to the worth of that end (or the effective means to that end).³⁰ A willing, for Kant, has both an *object*, the end that the agent wills to bring about, and a *determining ground*, an underlying value commitment to the worth or goodness (on certain conditions) of that end. The determining ground is the agent's subjective *rationale* for the adoption of her end. It is clear why every willing must have an end, since "without some end there can be no *will*".³¹ But why must every willing also have a determining ground? Humans, unlike other animals, have a power of choice, Kant maintains, that can "indeed be *affected* but not *determined* by impulses".³² Incentives usually only "tempt" us rather than determine us to act.³³ Of course, Kant does not claim that we are *never* simply determined by our impulses. Sometimes our emotions and desires are so powerful that they become what Kant calls "affects" and "passions"

which we momentarily cannot control,³⁴ in which case we are temporarily not fully responsible for our actions.³⁵ In such cases, and this includes examples of fits of rage (an "affect") and addictions (a "passion"), our will is effectively *bypassed*.

But we are not always simply determined by our impulses in these ways. An impulse which tempts us to act can be deferred or denied. In that case there arises a gap between our impulses and our adoption of ends. This gap exists because we can (at least sometimes) gain a reflective distance from our sensible impulses, and insofar as we can do that, we need to first-personally commit ourselves to taking (or not taking) the ends that our impulses prompt us to adopt as *worth* adopting. This involves a judgment that the proposed end is *good* and worth doing, and that judgment must involve an appraisal "by reason".³⁶ This commitment to the worth of that end is the determining ground of that choice. We do not need to, although sometimes we do, consciously make this commitment before we act. Sometimes we reveal our commitments, even to ourselves, by what we actually do. And often (or even always) we are not *certain* what the determining ground of our choice really is. Kant does not conceive of humans as, motivationally, perfectly self-transparent beings.³⁷

The job of the executive function of the will (*Willkür*) is to will ends and the job of the legislative function of the will (*Wille*) is to provide determining grounds.³⁸ Determining grounds must be provided by one's practical reason (or *Wille*), since it is reason that provides principles or rationales. But while "every volition must also have an object" or matter, that matter "is not just because of this, the determining ground and condition" of the choice.³⁹ Kant argues that "insofar as it [the will (*Wille*)] can determine choice [alone], it is ... practical reason itself".⁴⁰ This allows us to understand the difference between heteronomous and autonomous willing as being a matter of two different types of determining grounds. In both cases the determining ground of our choice must come from reason, since only reason can provide principles or rationales. In heteronomous willing it is "empirically conditioned reason", or the desire for the object of the will as a principle of empirical reason, that determines the will. In autonomous willing it is "reason itself" or "pure practical reason" that determines the will,⁴¹ that is, the autonomous principle of willing as if one were a member of a realm of ends by only adopting ends on condition that they are compatible with the dignity of all persons. In heteronomous

willing we make ourselves a slave to our desires, since those desires effectively give the law to our will, whereas in autonomous willing it is our own reason itself that sets the law that our desires must obey.⁴² Rational agency is therefore understood by Kant to be, not a battle of reasons and passions, or calm passions and violent passions (as it is for Hume), but of two different sorts of reasons or rationales (autonomous and heteronomous ones).

When we will autonomously the determining ground of the will in its legislative function must be pure practical reason itself. To this extent, the will in its legislative function cannot be at all influenced by any property of the object of the will if we are to will autonomously. However, all Kant's claims about autonomous willing relate to the *law* that determines the will, that is, to the primary determining ground of the will in its legislative function. But willing has both a law (or determining ground) and an object (or end). When we will autonomously, must the will in its executive function also not be at all influenced by any property of the object of the will? To answer this question we need to turn to Kant's hierarchal account of motivation in *Religion Within the Boundaries of Mere Reason*. There Kant says that we have a natural predisposition to care about our self-love or happiness and a predisposition to care about morality itself and that the will is always (or at least normally) influenced by both sorts of incentives. What matters, says Kant, is not the "material of the maxim" (or the object of the will), but whether agents will their happiness on condition of its compliance with the law of morality or will to act morally on condition of its compliance with the law of self-love. What matters is not the object of the will but which of the two laws of reason (empirical and pure) is the *primary* determining ground of the will; that is, "*which of the two* [his own happiness or the dignity of persons] *he makes the condition of the other*".⁴³ If both rationales are, as Kant says, always present, then the difference between autonomous and heteronomous willing must be in the order of one's commitment to these rationales.

What this implies is that when we will autonomously it need not be the case that our executive power of choosing (*Willkür*) is not at all influenced by the object of our will. Even when we will autonomously our desires may provide the impetus for us to consider adopting some particular end, and our desires are obviously influenced by external social and historical factors and the properties of the end or object we are adopting. Kant's conception of

autonomous willing, therefore, need not problematically require that we are not at all *influenced* by any external factors when we will an end, in the sense that our will in its executive function (in its choice of an object or end) must not be influenced at all by any property of that object or end. It need only require that I am not *determined* by them, that is, that the legislative faculty of my will (*Wille*) is not at all influenced by the object of my will when it legislates a *primary determining ground* (rather than an *object*). And it is not at all influenced by the object of my will when I place a *formal rational condition* (properly valuing the dignity of persons), rather than a *substantive empirical condition* (I happen to want this end), on the worth of choosing that end.

To see what this means in practice it will be helpful to investigate the hierarchical nature of commitments. This hierarchy is clearly revealed in cases of conflict. For example, it is my end to paint an inspiring landscape this weekend. However, my child suddenly falls ill on the weekend while I am painting and I need to drive her to see a Doctor. But this will prevent me from painting this weekend. Which commitment do I make the *condition* of the other? In this case I choose to drive my child to the Doctor. In making this choice I demonstrate that my commitment to painting this weekend is *conditioned* by a higher order commitment. More generally, our specific commitments to ends only make sense in terms of an underlying commitment to a personal conception of happiness or the good life.⁴⁴ But painting landscapes when my child is in need of urgent medical attention does not form part of my conception of happiness. Our specific commitments to the worth of ends therefore constitutes a hierarchy, with higher order commitments (the health of one's child) acting as conditions on the worth of lower order commitments (landscape painting), with a conception of happiness at the apex of this hierarchy of personal commitments.

But our commitment to our own happiness, and the sub-commitments that constitute it, need not be our *supreme* commitment. Our supreme commitment is the basic principle of our will which Kant calls our character-defining *disposition* (*Gesinnung*).⁴⁵ Our disposition defines our commitment to the unconditional worth of our highest order good which functions as the condition of the worth of all our lower order goods. Kant argues that there are only two candidate dispositional commitments, our own happiness (the heteronomous principle of self-

love) and the dignity of all rational beings (the autonomous principle of respecting dignity).⁴⁶ These two candidates are grounded in empirically conditioned reason (heteronomy) and pure practical reason (autonomy) respectively. When your commitment to the dignity of rational agents comes into conflict with your commitment to your own happiness, which do you make the condition of the other? When we will autonomously the *primary* determining ground of that choice is an unconditional commitment to the dignity of all persons. When we will heteronomously it is not. A perfectly autonomous person has an unconditional commitment to the dignity of all persons as his or her character-defining disposition. While one need not be a perfectly autonomous person to *on occasions* will autonomously, only a perfectly autonomous person will *always* will autonomously.

The commitment to the dignity of all persons, in its role as a determining ground of choice, can function either as what Barbara Herman calls a "primary motive" or a "limiting condition motive". The primary motive is the one that the agent acts on and that "can, by itself, produce an action". A limiting condition motive is also a motive that the agent acts on, but one that cannot by itself produce an action. Instead it acts as a limit on other motives.⁴⁷ An example of a limiting condition motive is the commitment to 'safety first'. First check that it is safe to do what you want to do, and only then do it. In the case of moral necessitation, where we must (or must not) act in a certain way, the principle of autonomy can function as a primary motive. In that case it says: dignity requires that you do (or not do) this. But in cases where an action is merely morally permissible, the principle of autonomy can function only as a limiting condition motive. In that case it says: first check that it is compatible with properly valuing the dignity of all persons to fulfil your conception of happiness, and only then fulfil it.

For example, I will autonomously when I will the end of landscape painting *on the formal rational condition* that doing so is compatible with respecting the dignity of all persons, that is, *on condition* that my willing this end is normatively authoritative for all other rational agents by introducing it into the realm of ends. A *subordinate* condition that I may put on my commitment to this end is that it *also* remains part of my conception of happiness. But my *primary* or *highest* commitment is to dignity first. If adopting this end becomes incompatible with that highest commitment, then I would no longer judge this end to be worth adopting. In contrast, I

will heteronomously when I will the end of landscape painting *on the empirical condition* that I continue to desire this end as part of my conception of happiness. In that case my commitment to my own happiness, and not a commitment to dignity, is the *highest* condition I put on the worth of my end. When I will heteronomously the object of my will also determines my will. When I will autonomously, pure practical reason determines my will, even when my choice of ends (but not my choice of primary rationales) is influenced by properties of the object of my will.

4. Autonomy Competences and Socialisation

Having outlined Kant's conception of autonomous willing, our next task is to situate that conception within the literature on personal autonomy. To do that we first need to examine, at least briefly, the connection between autonomy and socialisation in Kant's conception. This is a necessary first step since the connection between autonomy and socialisation, including the role of oppressive socialisation, is a central focus of contemporary conceptions of personal autonomy. If we are to situate Kant's conception within this literature then we need to know how Kant understands this connection. It is sometimes thought that for Kant there is no connection between autonomy and socialisation since our autonomy is completely independent of any external factors or forces. However, this simplistic interpretation fails to capture the nuance of Kant's view. Kant does claim that we each have a dignity or absolute worth, based in our *capacity* for autonomy (but not the perfect *realisation* of that capacity) that neither changes nor develops over time and is therefore independent of socialisation. However, he does not also claim, as we shall see below, that our efforts to perfectly realise our autonomy by willing autonomously are not at all influenced (even if they are not determined) by patterns of socialisation. It is this connection that we shall explore in this section.

According to Kant, our efforts to will autonomously are influenced (even if not determined) by the level of our achievement of various autonomy competences which are in turn influenced by historical patterns of socialisation and our individual self-development.⁴⁸ The autonomy competences required for the ability to be able to will one's ends on condition that one's will has normative authority for all rational agents are, according to Kant, certain

emotional propensities, as well as the abilities to reason, imagine, judge, and hold (at least temporarily) appropriate evaluative attitudes. While Kant thinks that we must assume that all morally responsible agents have these autonomy competences to at least a threshold level, each of these autonomy competences can be encouraged or hindered by patterns of socialisation.

These autonomy competences include moral feeling, conscience, love of human beings and respect, all of which Kant lists as predispositions on the part of feeling that must be assumed to exist in every human being if they are to be put under moral obligation.⁴⁹ But Kant is also clear that each of these feelings comes in degrees, and while it is our responsibility to cultivate each of them, it is also clear that Kant thinks that patterns of socialisation affect that task both positively and negatively. For example, Kant says that love of others will be hard to develop and maintain if you are frequently the victim of the ingratitude of others or if others never show love for you.⁵⁰ Alternatively, if others express gratitude for the love you show them, then this will tend to reinforce your love for others. Similar considerations apply to the other evaluative attitudes that morality, and thus the full realisation of our autonomy, requires that we develop and cultivate including, in particular, respect, love, and esteem for both ourselves and others.⁵¹

Self-respect is necessary since if I do not respect myself,⁵² then I will not think that my willing an end has normative authority even *for myself*, and if I do not think that, then I cannot regard my will as having normative authority for *all* rational beings. Rational self-love, that is "a predominant *benevolence* toward oneself" restricted to "the condition of agreement with" the moral law,⁵³ is necessary since if I do not care about myself, then I will not regard my own ends as valuable even *for me*, and I will therefore not regard them as important *for others*. Self-esteem is necessary since if I think of myself as worthless, then I will not regard myself as, and not expect others to regard me as, possessing dignity and a legislative will. For similar reasons, morality requires that we have practical attitudes of respect, (practical) love (that is, regard the happiness of others as important), and esteem toward others, since such attitudes are implicit in the demand that one wills one's ends on condition that they have normative authority for all rational agents. When one wills in that way, one expresses an attitude that the dignity and

personal ends of others are valuable in their own right and can limit the worth of one's own ends.

But each of these attitudes, just like the feelings that must be presupposed in a morally responsible agent, can be both cultivated and diminished and this, in turn, is influenced (but not determined) by patterns of socialisation. For example, Kant argues that it is through socialisation into norms of politeness in social intercourse that we first learn to treat others *as if* we owed them respect, love, and esteem.⁵⁴ And treating others in this way is an important first step toward actually *having* these attitudes. This emphasises that, in general, our evaluative attitudes towards ourselves and others are in part developed out of, and are reinforced by, patterns of intersubjective recognition.⁵⁵ As such, they can be influenced by negative and positive social forces. For example, if others do not treat you with respect, then it will be harder for you to maintain or fully develop a robust attitude of self-respect. Further, Kant's detailed accounts of individual moral (including childhood) development⁵⁶ and the connection he draws between moral, political and historical progress,⁵⁷ which we cannot explore in detail here (but which I have examined elsewhere),⁵⁸ add extra weight to the claim that for Kant, going through the proper process of development, which is partly dependent on historical social conditions, has an important influence on (even if it does not determine) one's evaluative attitudes and disposition and thus one's autonomy competences.⁵⁹

From this it follows that, on Kant's account, autonomy competences can be negatively influenced (even if not determined) by *oppressive* social conditions. (This helps to explain why Kant emphasises the moral importance of historical progress). By oppressive socialisation I shall mean here, a type of socialisation that reliably discourages or fails to properly develop the required autonomy competences in agents subject to that socialisation. This implies that we could think of Kant's conception of personal autonomy as belonging under the "umbrella" term of a *relational* theory of autonomy.⁶⁰ That is, a theory of autonomy premised on the claims that personal autonomy is encouraged by the right sort of socialisation and is discouraged by oppressive socialisation. However, while being subject to oppressive socialisation may make it more difficult for you to will autonomously, it does not make it impossible (or at least, not normally). In other words, for Kant failing to undergo the proper process of development does

not undermine one's moral dignity, or (except perhaps in very extreme cases) one's moral responsibility for one's actions, or one's ability to will lawfully (and thus to introduce ends into the realm of ends). This is because, on Kant's account: actions are morally "*imputed*" to persons who are held responsible for them insofar as they are *subject to* moral requirements and not insofar as they actually *fulfil* those requirements by willing autonomously;⁶¹ dignity is dependent on autonomy of the will and the capacity for autonomy and not on willing autonomously; and to be lawful an agent's will only needs to be *compatible* with the dignity of others.

On Kant's account historically dominant patterns of socialisations have tended to be oppressive. Kant sees the basis of this in the human predisposition to want one's worth as a person to be publically recognised by others. But this originally reasonable desire tends to get distorted by corrupt social conditions into "an unjust desire to acquire superiority for oneself over others".⁶² Such superiority is to be achieved by gaining more power, wealth, and status than the next person. Kant calls this desire for superiority over others our unsocial sociability. This unjust desire is based on the flawed view that the worth of persons is to be understood in terms of a person's relative possession of power, wealth, and status, rather than in terms of their equal dignity. For example, Paul Benson emphasises the way that feminine socialisation in Western societies tends to lead to the internalisation of the autonomy-discouraging view that a woman's *personal* worth is dependent on the status her physical appearance gives her.⁶³ This relative status is defined by how well, compared to other women, she does at meeting certain unrealistic standards of 'beauty'. Those who have a poor status in regard to this standard tend to lack self-esteem, self-respect, and self-love. However, men tend not to be less, but differently, oppressed, since men (and women) are often socialised into seeing a man's *personal* worth as dependent on, for example, how well he meets certain macho ideals, how much money he earns, or how conventionally 'beautiful' his wife is.

These specific (and contestable) examples are part of the more general problem that Kant identifies: we tend to be socialised into seeing the *worth of persons* in terms of their relative possession of positional goods (power, money and status), such that some persons are seen as having more worth *as persons* than others, rather than seeing the worth of all persons

as *absolute* and *incomparable* and grounded in their equal dignity. In oppressive social conditions the view that the worth of persons is determined by their relative status position is inculcated and enforced through socialising patterns of intersubjective recognition. This tends, Kant shows, to lead to widespread envy, ingratitude, and malice.⁶⁴ It leads to these attitudes, Kant argues, since someone who has internalised this oppressive view will tend to enviously regard others doing well as lowering their own worth, ungratefully regard the receiving of aid as lowering their own worth, and maliciously regard others doing poorly as increasing their own worth.⁶⁵ The internalisation of this oppressive view also leads agents to develop a conception of happiness that focuses on acquiring more power, wealth, and status than others. Governing yourself primarily in terms of such a conception of happiness and the associated practical attitudes of envy, ingratitude and malice, amounts to heteronomy, not personal autonomy.

However, while historical patterns of socialisation have tended to be oppressive, Kant also argues that, given the innate radical evil of human nature, it is likely that the full development of personal autonomy will only become *widespread* through the presence of widespread autonomy-promoting socialisation, which he optimistically hopes will come about through historical progress. Kant argues that: "however much the individual human being might do to escape from the dominion of this evil [of social corruption], he would still be held in incessant danger of relapsing into it ... the dominion of the good principle [autonomy] is not otherwise attainable ... than through the setting up and the diffusion of a society in accordance with, and for the sake of, the laws of virtues".⁶⁶ Without the social and institutional support provided by what Kant calls moral communities, that is, voluntary communities within a just political state committed to the promotion of virtue, Kant doubts that any (or at least many) humans could become perfectly autonomous persons (although, of course, we don't need to live in such a society to will autonomously on occasions or to be a less than perfectly autonomous person). Such moral communities would be premised on the equal dignity of all persons, and this constitutive value commitment would be institutionalised and reinforced through patterns of social interaction that express attitudes of respect, love, and esteem for all persons. Being socialised into such a community would help to foster, rather than discourage, the full development of autonomy competences in all persons. In summary, for Kant our personal

autonomy can be influenced (even if it not determined) both positively and negatively by social forces.

5. Kant and Contemporary Theories of Personal Autonomy

Having outlined Kant's theory of personal autonomy and examined the connection for Kant between socialisation and autonomy, we are now in a position to compare Kant's conception with contemporary theories of personal autonomy. Contemporary theories of personal autonomy can be classified as procedural or substantive.⁶⁷ Procedural theories take content-neutral procedures to provide necessary and sufficient conditions for autonomy. An example of a procedural theory would be one that asserted that *whatever* end I choose, provided I choose it on the basis of *whatever* practical identity or higher-order desires that I endorse through critical reflection, counts as an autonomous willing.⁶⁸ According to such theories the *content* of my willings and the *content* of my practical identity or higher-order desires are irrelevant to determining whether my willings are autonomous or not. However, procedural accounts tend to struggle at dealing with examples of oppressive socialisation. In that case a person may govern herself in terms of her own endorsed practical identity or higher-order desires, but she is still not autonomous since that identity and those desires are not *really* her own as they are the product of oppressive socialisation.⁶⁹

Substantive theories of autonomy attempt to deal with this problem by claiming that content-neutral procedural conditions are not necessary and sufficient for autonomy. As such, they set out substantive conditions that the content of willings, or the values that ground those willings, must meet in order to count as autonomous. A *strong* substantive theory is one whose conditions *fully* determine the content of *all* autonomous willings. A *weak* substantive theory is one whose conditions (unlike procedural theories) *somewhat* limit and inform the content of willings that can count as autonomous, but does not (unlike strong theories) *fully* determine the content of *all* autonomous willings. Weak substantive theories constitute a spectrum of theories that *more or less* limit and inform the content of autonomous willings.

According to strong substantive theories a willing can count as autonomous only if it has the right *content*, where a complete account of the right content can be given. An example of

such a theory would be one that asserted that only willings which involve the adoption of objectively good ends can count as autonomous, where a complete account of those objectively good ends is (or can be) given.⁷⁰ Strong substantive theories can handle the problem of oppressive socialisation since they are able to provide an objective standpoint against which an agent can assess and correct their values and ends. The problem with oppressive socialisation, on this view, is that it undermines an agent's ability to detect objectively good ends.⁷¹ However, strong substantive theories must rely on a *separate* account of objectively good ends and therefore need to explain why choosing such ends counts as *autonomy* or *self-government*. This leads to the further worry that strong substantive theories can't leave sufficient room for an agent to develop his *own distinctive* conception of the good.

An example of a weak substantive theory is one that asserts that whatever end a person wills, that willing can count as autonomous only if the person who wills that end has the evaluative attitudes of self-respect, self-love and self-esteem.⁷² The problem with oppressive socialisation, on this view, is that it undermines an agent's ability to hold these evaluative attitudes. Further, having these evaluative attitudes means that certain ends cannot be autonomously willed, such as ends which are incompatible with maintaining self-respect. But this still leaves plenty of room for an agent to develop her *own distinctive* conception of the good.

However, it is not clear that such a weak substantive theory can deal with the problem of oppressive socialisation in cases where those oppressive standards are actually met or exceeded. For example, Angela and Betty have internalised the (on Kant's account) oppressive norm that their worth as persons is dependent solely on their looking conventionally 'beautiful'. But Betty meets or exceeds those standards, whereas Angela does not. As a result, Betty has plenty of self-respect, self-love and self-esteem, whereas Angela has very little. This seems to imply, on this weak substantive account, that Betty but not Angela wills autonomously since she alone has the right evaluative attitudes. But intuitively that doesn't seem right. Both of them seem to be oppressed because both of them have internalised the *same* oppressive norm that their worth as persons is dependent solely on their looking 'beautiful'. The fact that Betty

happens to meet or exceed that oppressive norm doesn't seem to make her any more autonomous than Angela.

We are now in a position to answer the question that concerns us here: where (if anywhere) does Kant's conception of autonomous willing fit in this contemporary taxonomy? On the one hand, Kantian approaches to ethics and justice are often associated with procedural approaches. But Kant's conception of autonomous willing cannot be a procedural theory of autonomy, even if it has procedural elements, since it does not take the content of willings and values to be irrelevant. In particular, if the content of my willing is disrespectful of the dignity of any rational being, then my willing cannot be autonomous. On the other hand, Kantian approaches to autonomy have also been labelled as strongly substantive.⁷³ But this label misunderstands the role that dignity plays in Kant's conception. The role of dignity in the case of merely permissible willings, which are the central focus of accounts of personal autonomy, is that of a limiting condition (not a primary) motive. While the 'dignity first' clause does limit and inform a person's ends to some extent, it does not *fully* determine the *content* of those ends, since it leaves a person with sufficient scope to develop her own distinctive conception of the good on condition that it remains within those limits.

As such, we should classify Kant's conception as a weak substantive theory of personal autonomy. As with other weak substantive theories, Kant's conception is able to avoid the difficulties faced by procedural theories in dealing with oppressive socialisation by stressing that autonomous agents must obtain the required autonomy competences, and these include developing (or at least temporarily expressing) the right evaluative attitudes, such as self-respect. However, unlike other weak substantive theories, it stresses the importance for personal autonomy not only of *self*-attitudes, but also of *other*-attitudes. Kant's conception also stresses the importance, not only of having self-respect, self-love, and self-esteem, but of having these attitudes for the *right reasons*, namely because of the dignity that persons possess. This provides Kant's conception with a distinctive account of *when* an agent may *rightfully* regard his will as a legitimate source of normative authority for others, namely when his will is legislative in the realm of ends. Kant's conception also avoids the problems associated with strong substantive theories of autonomy by being able to explain why obeying a 'dignity

first' norm amounts to autonomy, namely, because it is a norm that is set by one's *own reason* itself. This makes it, overall, a promising theory of personal autonomy.

6. Further Elucidation of this View

On Kant's conception of autonomous willing, I will autonomously if I satisfy clauses A, B and C:

A) I must *will* an end.

B) I must will that end *on condition that my willing that end has normative authority for all rational agents*. This requires that:

B1) respect for the dignity of persons acts as either the *limiting condition* motive (for *permissible* ends) or the *primary motive* (for *obligatory* ends) for my willing of that end;

B2) I have the necessary *autonomy competences*.

C) My will must *really have* the normative authority for all rational agents that clause B requires that I *regard* it as having.

To illustrate these three clauses we shall look at four types of cases.

Case 1): I regard what I will as having normative authority for all rational agents and it has that authority. For example, James wills the permissible end (passes A) of landscape painting on condition that his willing that end has normative authority for others (passes B), and it has that authority (passes C). This counts as an example of autonomous willing.

Case 2): I do not regard what I will as having normative authority for all rational agents but it has that authority. This covers many cases of oppression. For example, Wilma has internalised, through socialisation into a sexist society, the oppressive norm that as a woman she is a second-class being who should be unconditionally submissive to the will of her husband. Wilma has thoroughly internalised this norm to the extent that she freely governs herself in accordance with it and feels at home in her sexist society where such norms are ubiquitous. But Wilma does not regard her willing of an end as having normative authority for all others persons (fails B) because she does not believe the will of a woman can have that sort of authority for a man. Wilma therefore does not will autonomously and she lacks self-respect.

However, although Wilma does not *regard* her will as legislative for others, her will *is* legislative for others when what she wills is compatible with the dignity of all persons. As such, her will can be *legislative* or have normative authority for others even when her will is not *autonomous*. Insofar as her will is legislative for others, her ends are still introduced into the realm of ends by her willing them and others still have reason to help her to achieve her (non-autonomously set) ends. Further, even though she does not will autonomously, Wilma still has dignity since dignity is grounded in being a rational being (and having autonomy of the will) and not in willing autonomously, and Wilma can still be held (at least partially) morally responsible for her actions since she is still *subject* to the principle of autonomy.

Case 3): I do not regard what I will as having normative authority for all rational agents and it doesn't have that authority. This covers, for example, many cases of wrongdoing. Since to will autonomously is to will an end on condition that you willing it has normative authority for all rational beings, it follows that it is impossible to autonomously will a morally wrong or evil action. To see why this is not an unintuitive implication consider the case of an "idealized Mafioso", Tony, whose practical identity is bound up with a "code of strength and honour" and who governs himself in accordance with this identity.⁷⁴ Tony, living up to his identity, wills the end of extorting money from Sam, a local shopkeeper. Is Tony willing autonomously in this case? No, since Tony (let's assume) uses Sam as a mere means. As such, his willing this end lacks normative authority for Sam (fails C). But Tony does not even *regard* his will as having normative authority for Sam (fails B). Insofar as what Tony wills has any authority over Sam, it has that authority not because Tony *willed* that end, but because Sam *fears* Tony. We can see this clearly if we imagine a slight modification to this example. Imagine that a policeman prevents Tony from extorting money from Sam and provides very creditable assurance (which both parties believe) that he will prevent any similar future extortion attempts. While Tony might be angry or annoyed by this, he would not demand that the policeman respect the normative authority of his will because he was not governing himself on that basis that others *should* regard his will as having that sort of authority. But then Tony is not willing autonomously. Instead he is allowing himself to be governed by his desire for money and for maintaining his menacing reputation.

Case 4): I regard what I will as having normative authority for all rational agents but it doesn't have that authority. This covers, for example, cases of mistaken or self-righteous wrongdoing. Imagine that it is not Tony, but his self-righteous brother Vincent, who makes it his end to extort money from Sam. As in the above example, a policeman prevents the extortion and any future occurrences. However, unlike Tony, Vincent is genuinely indignant about the policeman's interventions and demands that the policeman and Sam respect the normative authority of his will. Of course, Vincent's will lacks that authority as he is using Sam as a mere means (fails C), even though he (wrongly) *regards* his will as having that authority. As such, Vincent is not willing autonomously, even if he thinks he is.

Unlike Wilma, Vincent seems to have plenty of self-respect, self-love and self-esteem (although not much respect for others). But the problem is that Vincent respects himself for the wrong reasons, namely because he is a powerful and strong Mafioso, and not for the right reasons, namely because as a rational being he (like Sam and all other persons) has dignity. An agent who wills autonomously, unlike the self-righteous wrongdoer, does not regard what she wills as having normative authority for others *no matter what*. Rather, she regards her will as having that normative authority *on the condition* that what she wills does not infringe upon the dignity of others. This point allows us to return to the example of Angela and Betty given above. Recall, the fact that Angela had self-respect but Betty did not, did not seem to make Angela's willings autonomous and Betty's not autonomous since both had internalised the same oppressive norm. Kant's conception vindicates that intuition. Both Angela and Betty fail to respect themselves and others for the right reasons, namely because of the dignity or absolute worth that all persons possess. However, in their case (and arguably unlike in Vincent's case) this is partly due to the impact of the oppressive socialisation that they have been subject to.

Finally, it is worth noting that on Kant's conception it will not always be easy to know whether someone has willed autonomously or not. An agent wills an end autonomously if she wills *that end* on condition that it has normative authority for all rational agents and it really has that authority. This means that if a situation were to arise in which her commitment to that end were to become incompatible with the dignity of all persons, then she would renounce her commitment to that end in that situation. But that situation may never arise. In that case, can

she be *certain* that she *really would* renounce her commitment to that end *were* such a situation to arise? Probably not, but this just points to the general difficulty, which Kant emphasises,⁷⁵ of ever knowing for sure what the determining ground of one's choice *really* is. Such epistemological difficulties make it difficult to know for sure that one has willed autonomously. But this does not have the sort of problematic implications that it might have on other conceptions of personal autonomy. This is because on Kant's conception an agent does not need to will autonomously, or know for sure that he has or has not willed autonomously, in order to have a legislative will, or to have dignity, or to be held morally responsible for his actions. This reminds us that willing autonomously is a demanding ideal that we must strive for, not something that we should assume is always present.

7. Conclusion

Kant's conception of autonomy amounts to a *unified* theory of moral and personal autonomy, since you exercise your autonomy both when you do your moral duty on condition that respecting the dignity of others is your primary motive - call this *moral* autonomy - and when you adopt merely permissible ends in accordance with your own personal conception of happiness on the limiting condition that your will has normative authority for all rational agents - call this *personal* autonomy. This paper has therefore achieved its four goals by showing: that personal and moral autonomy need not be seen as in tension or at odds with one another; that Kant defends a promising weak substantive theory of personal autonomy; that socialisation can play an important role (both positive and negative) in the development of autonomy competences in Kant's theory; and that personal autonomy (properly understood) has an essential role to play in Kant's ethical framework.

¹ Thanks to Hugh Breakey, Catriona Mackenzie and to this journal's two anonymous referees for helpful suggestions and comments.

² Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon, 1986), 371; see also Joel Anderson and John Christman (eds.), *Autonomy and the Challenges to Liberalism* (Cambridge: Cambridge University Press, 2005), 2.

³ Marilyn Friedman, "Autonomy and the Split-Level Self," *Southern Journal of Philosophy* 24, no. 1 (1986): 20.

⁴ Paul Benson, "Autonomy and Oppressive Socialization," *Social Theory and Practice* XVII, no. 3 (1991).

⁵ Jeremy Waldron, "Moral Autonomy and Personal Autonomy," in *Autonomy and the Challenges to Liberalism*.

⁶ Immanuel Kant, "Critique of Practical Reason," in *Practical Philosophy* (Cambridge: Cambridge University Press, 1996), 5:33.

⁷ Robert Taylor, "Kantian Personal Autonomy," *Political Theory* 33, no. 5 (2005): 10-15.

⁸ Onora O'Neill, *Autonomy and Trust in Bioethics* (Cambridge: Cambridge University Press, 2002).

⁹ Kant, "Groundwork of the Metaphysics of Morals," 4:438 in *Practical Philosophy*.

¹⁰ Ibid., 4:440. See also Henry Allison, *Kant's Theory of Freedom* (Cambridge: Cambridge University Press, 1990), 94-5.

¹¹ Kant, "Groundwork," 4:432.

¹² Ibid., 4:438.

¹³ Onora O'Neill, *Constructions of Reason* (Cambridge: Cambridge University Press, 1989), 76.

¹⁴ Kant, "Groundwork," 4:441.

¹⁵ Kant, "Critique of Practical Reason," 5:33.

¹⁶ Allison, *Kant's Theory of Freedom*, 129-45, Kant, "The Metaphysics of Morals," 6:226 in *Practical Philosophy*.

¹⁷ Kant, "Critique of Practical Reason," 5:36.

¹⁸ Given our focus on autonomous willing, it will not be possible (or necessary) to engage with the extensive literature on Kant's account of autonomy of the will. But see, for example, Allison, *Kant's Theory of Freedom*, Karl Ameriks, *Kant and the Fate of Autonomy* (Cambridge: Cambridge University Press, 2000), Thomas Hill, *Autonomy and Self-Respect* (Cambridge: Cambridge University Press, 1991), Andrews Reath, *Agency & Autonomy in Kant's Moral Theory*

(Oxford: Clarendon, 2006), J B Schneewind, *The Invention of Autonomy* (Cambridge: Cambridge University Press, 1998).

¹⁹ Kant, "Groundwork," 4:433.

²⁰ Ibid.

²¹ It might be argued that personal (or private) ends are not part of the realm of ends for Kant since he says that he will "abstract" from the "content of their private ends" in order to "think of a whole of all ends in systematic connection" – Ibid. However, Kant does not *exclude* private ends from the realm of ends, but only *abstracts* from the *content* of those ends in order to *think* of the whole of all ends. Kant's point, then, is that we do not need to know the content of people's personal ends to think about the realm of ends as a whole. Even so, those personal ends, along with the ends-setters themselves, are part of the realm of ends.

²² Ibid., 4:433-34.

²³ Ibid., 4:436.

²⁴ Ibid., 4:434.

²⁵ An additional justification for this move is that Kant thinks that all the formulations of the categorical imperative are equivalent, although explaining how this could be so is a difficult task that is beyond the focus of this paper.

²⁶ This claim is not equivalent to similar sounding claims defended in Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), 106-33. The claim here is only about the normative status of personal ends in the realm of ends, and not about the source of goodness or value more generally.

²⁷ Kant, "Groundwork," 4:441. Similarly, in 4:440 Kant says: "Autonomy of the will is the property the will has of being a law to itself (independently of any property of the objects of volition)". See also 4:446-47.

²⁸ Ibid.

²⁹ Ibid., 4:444.

³⁰ Christine Korsgaard, *The Constitution of Agency* (Oxford: Oxford University Press, 2008), 57.

³¹ Kant, "On the Common Saying," 8:279-80 in *Practical Philosophy*.

³² Kant, "The Metaphysics of Morals," 6:213.

³³ Ibid., 6:380-81.

³⁴ Paul Formosa, "A Life without Affects and Passions: Kant on the Duty of Apathy," *Parrhesia: A Journal of Critical Philosophy* 13 (2011).

³⁵ Lewis White Beck, *A Commentary on Kant's 'Critique of Practical Reason'* (Chicago: University of Chicago Press, 1960), 33.

³⁶ Kant contrasts such judgments with the "mere feeling" that something is "agreeable" - Kant, "Critique of Practical Reason," 5:58.

³⁷ Allen Wood, *Kant's Ethical Thought* (Cambridge: Cambridge University Press, 1999), 202.

³⁸ Kant, "The Metaphysics of Morals," 6:213.

³⁹ Kant, "Critique of Practical Reason," 5:34.

⁴⁰ Kant, "The Metaphysics of Morals," 6:213.

⁴¹ Kant, "Critique of Practical Reason," 5:15. See also Andrews Reath, "Formal Principles and the Form of a Law," in *Kant's Critique of Practical Reason*, ed. Reath and Timmermann (Cambridge: Cambridge University Press, 2010).

⁴² Paul Guyer, *Kant's System of Nature and Freedom* (Oxford: Clarendon, 2005), 116.

⁴³ Kant, "Religion," 6:36.

⁴⁴ Kant, "Critique of Practical Reason," 5:34.

⁴⁵ Immanuel Kant, "Religion within the Boundaries of Mere Reason," in *Religion and Rational Theology* (Cambridge: Cambridge University Press, 1996), 6:25. See also Allison, *Kant's Theory of Freedom*, 136-45.

⁴⁶ Kant, "Critique of Practical Reason," 5:22.

⁴⁷ Barbara Herman, *The Practice of Moral Judgment* (Cambridge, Mass.: Harvard University Press, 1993), 1-44.

⁴⁸ On autonomy competences see Diana Meyers, "Personal Autonomy and the Paradox of Feminine Socialization," *Journal of Philosophy* 84, no. 11 (1987).

⁴⁹ Kant, "The Metaphysics of Morals," 6:399-403.

⁵⁰ Kant, "The Metaphysics of Morals," 6:455. See also Paul Formosa, "Kant on the Highest Moral-Physical Good: The Social Aspect of Kant's Moral Philosophy," *Kantian Review* 15, no. 1 (2010): 12-21.

⁵¹ As a matter of simplification I will sometimes speak as if these practical attitudes are all or nothing, rather than (as they in fact are) matters of degree.

⁵² Kant, "Critique of Practical Reason," 5:161, Kant, "The Metaphysics of Morals," 6:403. See also Eric Wilson, "Kantian Autonomy and the Moral Self," *The Review of Metaphysics* 62 (2008): 358.

⁵³ Kant, "Critique of Practical Reason," 5:37.

⁵⁴ Patrick Frierson, "The Moral Importance of Politeness in Kant's Anthropology," *Kantian Review* 9, no. 1 (2005).

⁵⁵ Joel Anderson and Axel Honneth, "Autonomy, Vulnerability, Recognition, and Justice," in *Autonomy and the Challenges to Liberalism*.

⁵⁶ Immanuel Kant, "Lectures on Pedagogy," in *Anthropology, History, and Education* (Cambridge: Cambridge University Press, 2007).

⁵⁷ Pauline Kleingeld, "Kant, History, and the Idea of Moral Development," *History of Philosophy Quarterly* 16, no. 1 (1999): 59-80.

⁵⁸ See Paul Formosa, "'All Politics Must Bend Its Knee before Right': Kant on the Relation of Morals to Politics," *Social Theory and Practice* 34, no. 2 (2008), Paul Formosa, "From Discipline to Autonomy: Kant's Theory of Moral Development," in *Kant and Education: Interpretations and Commentary*, ed. Klas Roth and Chris Surprenant (New York: Routledge, 2011), Paul Formosa, "The Role of Vulnerability in Kantian Ethics," in *Vulnerability: New Essays in Ethics and Feminist Philosophy*, ed. Catriona Mackenzie, Wendy Rogers, and Susan Dodds (Oxford: Oxford University Press, 2013).

⁵⁹ Does this lead to an element of "luck" in our moral evaluations? I shall not attempt to answer that complex question here but see, for example, Kleingeld, "Kant, History, and the Idea of Moral Development," 72.

⁶⁰ Catriona Mackenzie and Natalie Stoljar, "Autonomy Refigured," in *Relational Autonomy*, ed. Mackenzie and Stoljar (New York: Oxford University Press, 2000).

⁶¹ Kant, "The Metaphysics of Morals," 6:223.

⁶² Kant, "Religion," 6:27.

⁶³ Benson, "Autonomy and Oppressive Socialization."

⁶⁴ Kant, "Religion," 6:27.

⁶⁵ Kant, "The Metaphysics of Morals," 6:458-61.

⁶⁶ Kant, "Religion," 6:93-94.

⁶⁷ Mackenzie and Stoljar, "Autonomy Refigured," 13-21.

⁶⁸ Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *The Journal of Philosophy* 68, no. 1 (1971).

⁶⁹ Benson, "Autonomy and Oppressive Socialization", Mackenzie and Stoljar, "Autonomy Refigured." For a procedural attempt to deal with this problem see John Christman, "Autonomy: A Defense of the Split-Level Self," *The Southern Journal of Philosophy* 25, no. 3 (1987).

⁷⁰ Benson, "Autonomy and Oppressive Socialization".

⁷¹ Susan Wolf, "Sanity and the Metaphysics of Responsibility," in *Responsibility, Character, and the Emotions*, ed. Schoeman (Cambridge: Cambridge University Press, 1987).

⁷² Catriona Mackenzie, "Relational Autonomy, Normative Authority and Perfectionism," *Journal of Social Philosophy* 39, no. 4 (2008): 514.

⁷³ Mackenzie and Stoljar, "Autonomy Refigured," 20.

⁷⁴ Christine Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), 254-58.

⁷⁵ Kant, "Groundwork," 4:407.