

Norms of Truthfulness and Non-Deception in Kantian Ethics

Donald Wilson

Abstract: Questions about the morality of lying tend to be decided in a distinctive way early in discussions of Kant's view on the basis of readings of the false promising example in his *Groundwork of The metaphysics of morals*. The standard deception-as-interference model that emerges typically yields a very general and strong presumption against deception associated with a narrow and rigorous model subject to a range of problems. In this paper, I suggest an alternative account based on Kant's discussion of *self-deception* in the *Metaphysics of Morals*. I argue that we make the concern with respect for our capacity for inner freedom seen in the case of self-deception the model for deception in general. Focusing on the case of paternalistic lying, I claim that this approach yields a subtle and integrated account that promises the kind of resources we need if we are to be able to make headway with hard cases where deception may seem permissible.

Questions about the morality of lying tend to arise and be effectively decided in a distinctive way early in discussions of Kant's view. Commentators often begin with an analysis of Kant's false promising example offered in the effort to ground or illustrate a favored interpretation of the Categorical Imperative, and this analysis becomes the general model for duties prohibiting deception.¹ The details vary depending on the underlying interpretation, but accounts like this tend to converge on what Tamar Schapiro describes as a "deception-as-interference" model in which the central wrong in deception lies in the attempt to subvert or control another's agency and therefore in manipulation rather than simply dishonesty.² Viewed in this way, the respect for others' agency required in these accounts is understood formally in terms of concerns with their *capacity* to act under the auspices of their reason, resulting in the broad and very strong prohibitions on deception in general characteristic of this kind of account.

There are some familiar concerns with this account, perhaps most notably the troubling rigorism entailed by the idea of a formal and exceptionless duty to respect the purposive agency of others irrespective of their ends.³ More generally, it seems implausible to regard the deliberate

subversion of agency as a fundamental and very serious wrong without also being concerned with cases of accidental or incidental subversion of agency arising from carelessness or indifference in our dealings with others that may have exactly the same effects. I am not going to argue here that this common model is defective, but I think that problems like these suggest that there is room for an alternative account capable of distinguishing between different cases of deception and of connecting and integrating concerns with deception and honesty.

I argue here that the basic elements of such an account can be found in Kant's discussion of the duty prohibiting *self*-deception in the *Metaphysics of Morals*.⁴ The general idea is a simple one. It seems plausible to suggest that while there may be duties owed specifically only to oneself or others, duties prohibiting the same type of action ought to be related and share a similar moral concern.⁵ I propose, therefore, to approach the issue of lying to others from this perspective and to look to an understanding of the moral problem associated with *self*-deception for insight into concerns with deceiving others and, more generally, into the role of norms of truthfulness in a Kantian ethical view. Doing so, I argue, affords us a more comprehensive and integrated model with the resources we need to offer a more nuanced and plausible account of norms of non-deception and truthfulness.

I should emphasize at the outset, however, that my concern here is to outline an alternative account of norms of truthfulness that I think is suggested by aspects of the *Metaphysics of Morals* and specifically *not* to try to offer an account that is clearly in line with all of the various remarks that Kant makes about deception.⁶ My focus here is therefore on describing the basics of the alternative model I propose and contrasting it with the alternative deception-as-interference account. I begin with a discussion of the nature and scope of the duty prohibiting self-deception. In section III, I consider what this account suggests for our understanding of norms of

truthfulness and non-deception owed to others. Section IV compares the resulting account with the kind of deception-as-interference model offered by Korsgaard and others and argues that the model I propose has a number of advantages.

II

Kant classifies the duty prohibiting self-deception in the *Metaphysics of Morals* as one of a series of perfect duties of virtue owed to ourselves concerned with ensuring “inner freedom” (MM, 6:420) and thus “the capacity for self-constraint not by means of other inclinations but by pure practical reason” (MM, 6:396). This kind of freedom differs importantly from both the negative sense of freedom that Kant discusses – freedom from simple *determination* by sensible impulses – and from the positive notion of freedom associated with the legislative capacity of pure practical reason (MM, 6:213; CPrR, 5:33). As Stephen Engstrom⁷ notes, Kant’s earlier metaphysical account of practical freedom in terms of these notions directly implies the basic capacity for rational self-constraint constitutive of inner freedom.⁸ The notion of inner freedom adds, however, a distinctive element of *self-constraint* reflecting the specifically practical orientation of *The Metaphysics of Morals* and Kant’s concern with the application of formal moral principles to the specific circumstances of human agency. Human beings are imperfectly rational beings subject to other incentives for whom rational self-government requires conscious self-constraint and effort in the face of competing incentives. Moreover, this susceptibility of our wills to influence by sensible affection is not a contingent aspect of our agency: We *always* remain subject to other incentives and *cannot* simply become purely rational agents for whom the deliberative authority of principles of pure practical reason is a given. Beings like us therefore realize our rational nature *only* in the conscious choice of rational self-government and hence *only* in the development and affirmative exercise of our capacity for inner freedom.

Concerned with ensuring our capacity for this kind of inner freedom, the perfect duties of virtue that Kant discusses here are therefore concerned with preserving our basic capacity for moral agency. In the case of the duty prohibiting self-deception, this concern takes the form of a distinctive and deep worry about the integrity of moral deliberation. Kant thinks of moral deliberation as a kind of conversation with our conscience in which we articulate to ourselves a conception of what we propose to do and why we intend to do it and submit this account to the scrutiny of our conscience as our “inner judge” (MM, 6:400-401). Effective moral reflection and deliberation therefore depend essentially on the veracity and accuracy of the “evidence” submitted to the court of conscience, and the particular danger of self-deception lies in the way it insidiously compromises this process by bringing to it a distorted account of what we do and what we are like.⁹ Thus understood, self-deception is manifested in a kind of insincerity or lack of conscientiousness in reflecting on one’s actions and character that causes us to fail to impartially and honestly consider our choices and character and, instead, to allow our desires to dictate the understanding of our conduct we submit to our conscience for judgment.

Viewed in this way, self-deception consists in a failure to pay proper attention to what we are doing and what we are like that is regarded as a particularly troubling and insidious vice: one that we may fall prey to without even realizing it and one that undermines our capacity for inner freedom at the deepest level by compromising the deliberative processes on which this freedom essentially depends. Kant regards us as especially given to this vice. He also thinks that at least some lies told to *others* are violations of the duty prohibiting lying owed to *ourselves*. As I interpret this claim, this is because he thinks that self-deception tends to extend itself “outward” into our dealings with others, compounding the corrosive influence of desire on self-understanding. The opinions of our friends and those in our wider communities are an important

source of self-knowledge. Being honest with others affords us a fresh perspective on our actions, allowing us to see ourselves as others see us and thus to gain a deeper understanding of our actions and character. Conversely, deliberately or inadvertently and carelessly misrepresenting ourselves in communicating our thoughts to others compromises this important source of self-knowledge, compounding our vulnerability to self-deception by allowing us to avoid the kind of public accountability that can force us to recognize and take responsibility for our choices and character. Lying to others can therefore facilitate self-deception and Kant thinks of lies like this as compounding our tendency toward self-deception and contributing to a kind of systematic moral complacency to which he thinks we are naturally prone. In his *Religion Within the Bounds of Mere Reason* he stresses that we are naturally powerfully influenced by desire and tend to deceive ourselves with shallow and uncritical self-scrutiny, pursuing our desires giving little or no real thought to the moral character of our actions. At the same time, however, he thinks that we nonetheless tend to act well in public, getting along with others and conducting ourselves civilly in order to avoid disapproval and other social and legal sanctions that will make our lives more difficult. In this way:

This dishonesty, by which we throw dust in our own eyes and which hinders the establishment in us of a genuine moral disposition, then extends itself also externally, to falsity or deception of others.

Thinking too little about our character and our actions, and concentrating only on getting what we want while avoiding discord and conflict, we thus go through life deceiving others just as we deceive ourselves, and, basking in the approval of our peers and communities, complacently believe we are good (RR 6:38-9).¹⁰

Guarding against the vice of self-deception therefore requires more than merely refraining from the vice. Given our powerful tendencies towards special pleading on behalf of our interests

and to superficial self-evaluation, we risk falling prey to this vice without being aware of it and can resist it only through a robust and positive *active* commitment to honesty in self-reflection and in our general dealings with other people. Kant also, however, emphasizes the psychological dangers of obsessive self-reflection and practical concerns associated with excessive candor. He stresses, for example, that we are not responsible for all the transient thoughts and ideas that may occur to us “unbidden and spontaneously” and suggests that “eavesdropping” on ourselves by “occupying ourselves with spying out the involuntary course of our thoughts and feelings” is a sure route only to mental illness (PA, 7:133-134). Likewise, he stresses the appropriate limits of candor in our dealings with others, emphasizing that exposing our faults to others leaves us vulnerable to exploitation and may make them more liable to acquire our failings from our example (MM, 6:472). Elsewhere, he emphasizes the importance of frank and open discourse to the “health of our understanding” (PA, 7:219) and, contrasting casual acquaintance with deeper friendships, stresses that one of the important benefits of the latter lies in the way that close friendships facilitate a better understanding of controversial issues by allowing us to frankly exchange ideas and evaluate our opinions in the light of others’ views in ways that would be too risky in more casual relationships (MM, 6:472).

What seems to be required of us here is therefore a serious and conscientious effort at ongoing and careful self-scrutiny and a positive commitment to representing ourselves honestly in self-reflection and in our dealings with others, but one that is *informed by consideration of the nature and circumstances of our agency*. The concern here is one with respect for the conditions under which rational self-constraint is possible for human agents whose deliberative capacities are hostage to this kind of covert influence by desire, who are limited in their capacity to understand themselves and their choices, and who are exposed to other risks in the effort to do

so, and these and other factors relating to the psychological dynamics of our agency will affect the way the duty of truthfulness owed to ourselves is understood.¹¹

III

My suggestion in this paper is that we can use this account of the moral concern at issue in *self*-deception to inform our understanding of deception in general. I propose that we think of duties requiring us not to deceive others as sharing the same basic concern with inner freedom and thus as requiring a similar respect for the capacity for rational self-constraint *in others* and for the deliberative processes on which this capacity depends. To do so, however, we have to begin by recognizing that concerns with others' inner freedom can only take a distinct and necessarily more indirect form. We do not realize our inner freedom in acting in conformity with rational principles merely coincidentally or instrumentally in the pursuit of some other goal. We realize this freedom (and hence our rational nature) only in conscious *self*-constraint expressing *our own* internal ordering of values and priorities and a *reflectively endorsed* and *general* commitment to prioritizing the requirements of reason over our other interests. In this sense, realizing the capacity for inner freedom and preserving and strengthening it over time is *strictly* a personal affair – I can make it harder or easier for someone to come to make this kind of commitment but I *cannot* make it for them.¹² As such, a concern with others' inner freedom can therefore *only* take a more indirect form. To the extent that we think of this freedom as realized in the authentic expression of the agent's personal values and priorities, a concern with others' inner freedom requires a basic and robust respect for their *general* freedom to order and control their own lives on the basis of their own priorities and with their capacity to come, through doing so, to realize their moral natures for themselves.

These concerns will again require consideration of material and social circumstances and psychological factors, in this case, those bearing on the capacity of others to order and control their lives in general and to realize their moral nature in this way. These will include both concerns with adverse conditions incompatible with the development and realization of inner freedom and with those positively conducive to it. So, for example, the required concerns will commit us to basic norms of non-interference and to the kind of social world respectful of others' external freedom that Kant describes in the Doctrine of Right. More positively, however, these concerns will also include a broader range of norms enjoining an active sensitivity to the efforts of others to organize and control their own lives and that we seek to ensure a public culture respectful of different choices and supportive of others' efforts to realize their moral natures for themselves.¹³

Thus understood, the inner freedom model will involve a range of basic and robust concerns with what is referred to by Joseph Raz and others as "personal autonomy." According to Raz, "The ideal of personal autonomy is the vision of people controlling, to some degree, their own destiny, fashioning it through successive decisions throughout their lives."¹⁴ On the analysis I propose, however, this kind of personal autonomy is valued derivatively, though seriously, as a necessary condition of inner freedom and, crucially, the respect required for it is therefore *bounded and informed by this broader context*. In particular, the underlying requirement of respect for the conditions necessary for inner freedom gives us no reason to support individuals in choices that will undermine the realization of this capacity in themselves or others and this will allow us to distinguish between the ends agents seek in deciding what is owed to them.

Duties of non-deception owed to others on this basis will therefore take a very different form. Like the concern with self-deception, our understanding of the relevant duties here will again

encompass a broader and more positive range of commitments to norms of truth-telling: one that will require us to avoid not just deliberate manipulation but also carelessness and indifference in our dealings with others. If you ask me directions to get to an important job interview I might lie because I'm also a candidate and want to eliminate my competition. Equally, however, I may carelessly misinform you or only partially respond to your question just to save myself time in going about my own (unrelated) business or simply because I cannot be bothered to take the trouble to properly respond. These kinds of actions are also incompatible with the required concern with the capacity of others to order and control their own lives and, while we can certainly think of deliberate manipulation as worse, will therefore be regarded as wrong on the same general grounds. More positively, a requirement of respect for the development and realization of others' inner freedom like this will also encompass an active sensitivity to their deliberative processes enjoining us, for example, to volunteer information we have reason to think would be relevant but which may not have been actually requested and to support social efforts to make important information available in easily accessible ways (public libraries, low-cost internet access, etc.).

Like the obligations requiring us to respect our own capacities for inner freedom, the specifics of obligations of truthfulness owed to others will again reflect facts about our nature and circumstances and limits entailed by the rationale underlying these obligations. In acting in ways that respect others' capacities for inner freedom, we should be generally respectful of their efforts at organizing and pursuing their own lives and goals. To this end, we should avoid deliberately or inadvertently misleading others by word or by deed and should try to actively support them in their endeavors by providing needed information. Honoring these commitments will require us to take into account practical limits on our own and others' time, resources, and

understanding, and will sometimes require careful thought about what sort of information is appropriate in particular cases and how best to communicate it. So, for example, if I know that the important information needed by someone is somewhat technical and that they may lack the expertise necessary to make good use of it, I ought to take more time in explaining it and try to do so in ways that are likely to facilitate effective communication. Likewise, if I know that someone is unlikely to process information well in some case, because, say, they are in pain or under stress, I should take this into account and either tell them later or, if the information is needed now, make a special effort to avoid any misunderstandings on their part.

However, while we should be generally responsive to the information needs of others and be willing to help them when we can, we also have to be mindful of differences in the nature and relative importance of information needs and of other demands on our time and efforts. Decisions we make here should reflect the important basic concern with others' general freedom to determine their own lives and this will give us reason to prioritize interests central to the agent's sense of herself and her organizing commitments over requests for information about how to satisfy passing whims or occasional desires.¹⁵ In doing so, however, we should be wary of generalizations and should not think of ourselves as applying formal moral rules requiring a determinate order of priorities. The focus here is on respect for the capacity of others to determine the course of their own lives and, while we can sometimes rely on generalizations about the normal priorities people will have, different individuals will have very different views about how they want to live and what they value and we should be careful to avoid simply imposing our priorities on them in the judgments we make in these cases.

Importantly, obligations of truthfulness owed to others are also *limited* by this underlying concern with respect for our own and others' inner freedom. We should not discount others'

information needs simply on the basis of our own personal judgments about the value of their ends and goals, but, equally, we are not obligated to provide information that we have good reason to think is incompatible with the required respect for their own or others' capacities to realize and express this innate freedom and do not therefore have to provide information that will facilitate them in undermining their own or others' rational agency. Thus, I am not obligated to facilitate self-destructive behaviors in others and am free, instead, to refuse requests for information and to criticize others' behavior and try to persuade them to modify it for the better. Nor, similarly, does a requirement that I respect the general freedom of others to live their own lives grounded in the idea that it is only through doing so that they can come to realize their innate capacity for inner freedom, obligate me to aid others in immoral behavior that affects others. Individuals should be free to decide for themselves how they are to live and to make their own mistakes, but in doing so they have no claim to the freedom to choose to harm others and I have no obligation to aid them in doing so.

Given the need to use judgment here in assessing the facts of cases, we can expect to see reasonable disagreement as individuals make different judgments in particular cases about the relative priority of information needs and other considerations. In addition, while the specific content of our obligations will be partially fixed by facts about our nature, we can also expect to see a degree of cultural variation in the way these and other obligations are understood. So for example, to the extent that it is true that our psychological health depends on our being able to reserve aspects of ourselves we reveal only to friends and loved ones, we should expect this fact to be reflected in limits on norms of truthfulness and, more generally, in broader norms of privacy in a moral culture respectful of inner freedom. Within the limits set by this and other similar concerns, however, the specifics of our obligations will also reflect aspects of the history

and circumstances of particular communities and the need to balance concerns with privacy against other norms, and we can therefore expect to see variation across different moral communities. There is therefore no reason to expect our obligations here to be precisely codified in universally applicable rules meant to govern completely in any and all circumstances and no reason to think it would be helpful to try to do so. Instead, we should think of ourselves as subject to general rules that set out the basic requirements of respect for our own and others' inner freedom and function like Herman's deliberative presumptions in serving to guide our ordinary conduct and to inform the terms of occasional moral deliberation.¹⁶ In the case of norms of truthfulness, these rules will identify the different expectations of truthfulness ordinarily regarded as appropriate in different contexts and circumstances in a particular moral community and will typically suffice to guide our conduct. In more unusual cases where different presumptions seem to conflict, we will need to exercise judgment and, using these presumptions as guides, should reflect on the rationale underlying them and on the norms appropriate for a community of rational agents committed to the realization of their own and others' innate capacities for inner freedom in circumstances like these.

Viewed in this way, the specifics of obligations oriented around concerns with our own and others' inner freedom will routinely depend on social circumstances and the details of cases, and on a broad range of facts about the internal and interpersonal dynamics of human agency bearing on the integrity of deliberative processes essential to, or constitutive of, the capacity for inner freedom. This is not, however, to think of these duties as oriented around the goal of ensuring or maximizing the realization of our own or others' inner freedom. The obligations here are principled ones based on the formal moral requirement of respect for rational nature as this applies in the case of human agents who realize their rational nature only through the

development and exercise of their capacity for inner freedom. Thinking of a realized capacity for inner freedom as the empirical form of rational nature in human beings, the suggestion is that in our case the formal moral requirement of respect for the unconditioned value of rational agency takes the form of a requirement of respect for the conditions under which this freedom is possible. To understand what this requires, we must take into account various facts about us that bear on our own and others' capacities for inner freedom, including facts about the limits of human understanding; the form that moral deliberation and rational self-constraint take in human beings; and the various ways in which our deliberative capacities can be influenced and compromised. This does not commit us to thinking of a concern with this kind of freedom as a goal to be promoted. Nor is there any suggestion that we are entitled to violate norms of honesty where there is unlikely to be any actual harm of the relevant sort, or where we have reason to think some greater good may be served. Rather, the claim is that when properly understood, the relevant principled concerns may not apply in some situations and, when they do apply, may do so in different ways in different circumstances.

IV

Thus understood, the inner freedom model of norms of non-deception differs significantly from the standard account. The standard model favored by Korsgaard and others emphasizes norms of non-manipulation grounded in a formal requirement of respect for the capacity for autonomy definitive of our rational nature. Adopting a "thin" conception at least of the fundamental commitments of respect for autonomy, this account emphasizes the independence of our wills from external determination and then takes the requirement of respect to be one obligating us to fundamental norms of non-interference. The wrong in deception then consists in manipulation rather than falsehood, and deception is understood to be a central wrong simply in

virtue of its form, yielding the characteristic broad and serious prohibitions that are a distinctive feature of this account. Any more positive commitments to truth-telling are then viewed as distinct requirements associated, presumably, with obligations of kindness owed to others typically viewed as less serious.

In contrast, the inner freedom account understands the basic requirement of respect more broadly as one requiring respect for our own and others' capacities for inner freedom. Since this freedom is realized only in an enduring state of character defined by a substantive and self-conscious commitment to a distinct set of priorities, respect here requires a basic and important concern with the general agency of others and their freedom to determine their own values and priorities. Manipulation is again regarded as a central wrong because of the attempt it embodies to usurp the deliberative authority of another, but the concern in this case is regarded as continuous both with a broader range of concerns with truthfulness and with the wider conditions under which human beings are free to decide for themselves how to live and to come to realize their moral nature for themselves. To the extent that it involves a deliberate disregard for others' ends and goals and a willingness to usurp their control for one's own purposes, manipulation will still be regarded as a particularly serious failure of respect for others. To the extent, however, that careless disregard or deliberate neglect of the information needs of others is similarly incompatible with others' efforts to order and control their own lives, manipulation is viewed as continuous with failings like these and not as a special wrong in a distinct category. Thus, while it would be wrong for me to deceive someone in order to secure a better position in the line for movie tickets, it would also be wrong here for me to fail to provide needed information about an important matter to someone because doing so might cause me to be late for the movie¹⁷.

Thus understood, norms of truthfulness and non-deception require an active sensitivity to the

effect of a broad range of internal impediments to rational willing. The standard account focuses on cases involving deception but will presumably also be concerned with internal impediments to choice that may be exploited by others intent on manipulating individuals for their own purposes. So, for example, I presumably also ought to refrain from manipulating someone by exploiting my knowledge of his fears or passions in order to get him to act as I desire. In contrast, while the inner freedom account will likewise require us to refrain from exploiting these internal impediments to choice, it will also require that we actively take them into account in our dealings with others and try to avoid inadvertently or carelessly causing others to act hastily and without due deliberation on this kind of basis. In addition, because it treats norms of non-deception as derivative from broader concerns with truthfulness and thinks of these broader norms, in turn, as oriented around concerns with respect for the necessary conditions of our own and others' inner freedom, this account allows us to distinguish between the interests of individuals at stake in different cases and to prioritize ends and interests in terms of this underlying concern. By doing so, and by requiring a more positive sensitivity to the internal disposition of agents' wills, the inner freedom account affords us a more promising analysis of a range of cases.

To illustrate this, I'm going to end with some brief remarks on the different handling of cases of paternalistic lying in these models. In the case of mature adults with normal capacities, the standard model makes no distinction between paternalistic deception and deception motivated by self-interest with both being regarded as equally wrong simply in virtue of their deceptive form. As a result, the standard model rejects any paternalistic interference with others, permitting this only in cases where the capacity for autonomy can be said to have been temporarily or permanently lost. A very strong repudiation of paternalism like this is generally in line with the

widely shared moral intuition that people should be free to live their own lives.¹⁸ The strength of the prohibition in this case comes, however, at some cost. Given its emphasis on the bare form of deceptive acts, the standard model is unable to distinguish between cases that seem importantly different from the moral point of view and generally unable to capture our sense that the ends involved may complicate and sometimes change the moral character of acts of deception. Secondly, despite our deep general reservations about paternalism, some cases where deception is the only way to avert a very serious harm to the victim seem warranted. Given its blanket repudiation of deception, however, the standard model is forced either to prohibit paternalistic lying even in these cases, or to treat lies like these as cases of permissible non-deceptive falsehoods told to beings who lack autonomy. As a result, it either refuses to even consider countenancing lies that seem intuitively permissible, or insists on reducing the victim of the lie to the status of a child in order to permit them.¹⁹

In contrast, while the inner freedom account will also take prohibitions on paternalistic lying very seriously, it affords us a more nuanced account of the relevant norms allowing us to distinguish different cases and can be viewed as compatible with a limited permission to deceive. On the inner freedom account, obligations of truthfulness and non-deception require us to respect the conditions necessary if others are to be free to determine for themselves how they are to live and to come, through doing so, to realize their own innate capacities for rational self-constraint. This kind of self-determination and personal development is a process: we do not start out our lives capable of effective self-management with a mature and comprehensive conception of our own good, we come to this through a process of education and development that inevitably involves making mistakes along the way that deepen our understanding and strengthen our character. A commitment to respect for peoples' efforts to live their lives on their own terms will

therefore ordinarily protect others' liberties to make even what could reasonably be judged to be relatively serious mistakes in deciding how to live.²⁰

However, while the inner freedom account will therefore naturally encompass the same kind of robust and general prohibition on paternalistic interference thought of as the strength of the standard model, it will not protect all choices that others make in this way. The respect required of us for others' attempts to live their own lives obligates us to be generally frank and truthful but does not encompass choices that will clearly undermine deliberative capacities essential to other's agency and their capacity to realize their moral nature for themselves. Given the importance of learning mistakes here, our first response in cases like these should be to refuse to provide information that will contribute to these harms and explain why we will not do so. There may, however, be cases in which the harm involved is a serious one and refusal will not suffice to prevent it (because, say, the person seeking the information will take refusal or silence to confirm a particular view of things) and, in these narrow circumstances, active deception may be viewed as compatible with the requirement of respect for inner freedom and hence as a permissible last resort.

Imagine, for example, that your friend Mary has recently extricated herself from an abusive relationship with John.²¹ Assume that although Mary possesses the capacity for autonomy, she is nonetheless still suffering from the psychological effects of this abuse and from the loss of a significant long term relationship central to her life, and that she continues to have great difficulty in trying to free herself from John's influence and begin to live her own life.²² Imagine now that Mary is wavering in her resolve and clearly inclined to try to rekindle the relationship, and that she asks you directly whether you think John will be willing to do so. You know that he will readily take Mary back and have every reason to believe that the abuse will continue and is

likely to result in permanent and serious harm to Mary's psychological health and to her capacity to determine the course of her own life. Under these conditions, the normally robust protection afforded to learning mistakes is set aside and the requirement that I act in ways that express respect for Mary's capacity to lead a self-directed life and to realize her moral nature for herself gives me no reason to protect choices like these or to provide information that may facilitate them. Given normal limits on our knowledge of situations and outcomes, the strong presumption in favor of allowing people to make their own mistakes, and the special obligations of trust and good faith associated with friendship, our first response even in cases like this should still be to decline to answer and explain our position. Imagine, however, that we have very good reason to think that Mary will simply take our refusal to answer her questions as confirmation of her hopes. In this circumstance, knowing that silence will only be taken to confirm one understanding of the situation and will thereby predictably contribute to these harms, the inner freedom account can be regarded as compatible with deception or dissembling intended to divert Mary from her course.

Any permission to deceive here will, however, be tightly limited and constrained by the overarching requirement of respect for the conditions under which rational agency is possible and can flourish. It will not be enough merely to claim that a choice someone is about to make may impair their ability to reason well in the future. We would first need to be confident in our judgment of the facts of cases and wary of the influence of our own interests on our decisions. In many cases, deceiving others will be an easier course than confrontation and discussion. In the case of loved ones, it may also be a means by which we can spare ourselves the discomfort and pain of watching them harm themselves, and we will need to be sure that reasons like these do not cloud our judgment.²³ Even if these various practical considerations are met and we have

good reason to think that the agent's choice is likely to injure them in some way, the normal protection afforded to mistakes and bad judgment still applies and we would need some additional reason to think that the choice in question is a special case, so seriously self-injuring as to transcend this normally robust protection and justify an exceptional response. Even then, given the broad requirement of respect for others' agency and the risk of mistakes in our judgment about cases, our first response should still always be to decline to provide needed information and make our reservations explicit (thereby allowing the affected party an opportunity to respond), and we should consider deception only when we have some special reason to think that refusal will not suffice.

Lastly, even in cases where we have good reason to think that the usual norms of truthfulness may not apply, our actions must still be guided by the overarching requirement here enjoining respect for others' inner freedom and remain answerable to this requirement. In particular, if we are serious in thinking of deception as expressing respect for another's agency, we must also see it as part of a broader and more comprehensive response intended to support the victim in their efforts to direct their own life and we should therefore be committed, when possible, to appropriate longer term solutions to the underlying problem that prompts deception in the first place. We cannot, for example, hope to help Mary in dealing with being a victim of abuse merely by preventing her from rekindling her relationship with John on one occasion. Nor can we do so by systematically deceiving her in an effort to shield her from John, or other individuals like him, in the longer term. The real solution to the problem Mary faces consists in her confronting and overcoming whatever underlying problems prompt her self-destructive behavior in the first place and, if we are sincere in thinking that respect for her agency warrants deception in this case, we ought to also be committed to broader and more comprehensive solutions intended to help her

constructively and should think of any permission to deceive as being extended only under these conditions.

V

The issues here are complex and space is limited, but I hope that enough has been said to suggest that this account is worth considering further. The inner freedom account preserves a generally very strong presumption against paternalistic interference without committing us to the kind of absolute prohibition associated with the standard model. The respect required of us for others' attempts to live their own lives obligates us to be generally frank and truthful but does not protect choices that will clearly undermine deliberative capacities essential to other's agency and their capacity to realize their moral nature for themselves. If asked to provide needed information in these cases our first response should be to refuse and explain why we will not do so, but in cases where the harm involved is a serious one and refusal will not suffice to prevent it, active deception may be permissible as a last resort. We can certainly anticipate disagreement about whether reservation and deception may be appropriate in particular cases but there is an obvious moral difference between deceiving someone to benefit oneself or to induce them to act as you think they should, and doing so to save them from some serious self-inflicted harm. We need an account that acknowledges our deep sense of ambivalence in the latter kind of case and that gives us some constructive way of engaging the issues these cases raise without our having to infantilize the victims of paternalistic lies. I think that the inner freedom account promises a more nuanced analysis with the resources to do so.

We also need an account that will allow us to make similar distinctions in the case of lies told to benefit third parties or to save them from harm. I cannot pursue these issues here but I think the inner freedom account promises the same subtlety in relation to these cases. As we have seen,

this account begins with an active commitment to honesty in our dealings with others that requires a positive sensitivity to the information needs of others: one that is informed by the underlying rationale here and takes into account practical limits on our knowledge and time and other relevant considerations relating to the realization of our capacity for inner freedom, including psychological concerns with our need for intimacy and pragmatic concerns with our vulnerability to exploitation by others. Requiring a central respect for others' capacities to determine the course of their own lives, these norms will emphasize the importance of core commitments, allowing us to distinguish between these and less serious interests and concerns and to prioritize the former in choices about how to help others. In a mature moral community, these and other concerns will be reflected in the adoption of social norms relating to truth-telling distinguishing contexts of discourse in which mutually understood reticence is assumed from those in which truthfulness is more rigorously expected and will allow us to prioritize different information needs based on the importance agent's attach to different concerns.

These norms will robustly protect the efforts of individuals to organize and lead their own lives and will therefore support the same kind of strong deliberative presumption against deception endorsed by Barbara Herman and other standard models. Importantly, however, the underlying rationale also limits the scope of our obligations. I am obligated here to respect others' choices of ends and to try to provide them with needed information on the grounds that the general freedom to live as they choose is essential if they are to come themselves to the kind of distinctive choice of values and priorities in which their capacities for inner freedom are uniquely realized. As we saw, however, this background commitment gives us no reason to support choices that are obviously incompatible with the realization of inner freedom in the agent herself or in other people, allowing us to distinguish requests for information in cases like this

and treat them differently. This suggests that the inner freedom account promises a similarly nuanced approach to deception intended to benefit third parties; one that takes into account the ends sought in cases of deception and that preserves a generally very strong prohibition on deception without committing us to the unforgiving rigorism of more standard models. By doing so, and by affording us an integrated account of norms of truthfulness and non-deception, the inner freedom account promises the kind of comprehensive and nuanced account that Kantian ethics requires.

Bibliography

- Bok, Sissela. *Lying: Moral Choice in Public and Private Life*. Knopf Doubleday Publishing Group, 1999.
- Engstrom, Stephen. "The Inner Freedom of Virtue," *Kant's Metaphysics of Morals: Interpretive Essays*, Mark Timmons (Ed.). Oxford: Oxford University Press, 2002.
- Gregor, Mary J. *The Laws of Freedom*. New York: Barnes & Noble, 1963.
- Herman, Barbara. *The Practice of Moral Judgment*. Massachusetts: Harvard University Press, 1996.
- _____. *Moral Literacy*. Cambridge, MA: Harvard University Press, 2007
- Hill, Jr., Thomas E. *Autonomy and Self-Respect*. Cambridge: Cambridge University Press, 1991.
- _____. *Dignity and Practical Reason in Kant's Moral Theory*. Ithaca, NY: Cornell University Press, 1992.
- _____. *Respect, Pluralism, and Justice*. Oxford, Oxford University Press, 2000
- Kant, Immanuel. "Critique of Practical Reason," *Practical Philosophy*, trans. Mary J. Gregor & Allen Wood. Cambridge: Cambridge University Press, 1996. Abbreviated as CPR.
- _____. "Groundwork of The metaphysics of morals," *Practical Philosophy*, trans. Mary J. Gregor. Cambridge: Cambridge University Press, 1996.
- _____. *Lectures on Ethics*, trans. Peter Heath. Cambridge: Cambridge University Press, 1997.
- _____. *The Metaphysics of Morals*, trans. Mary J. Gregor. Cambridge: Cambridge University Press, 1991. Abbreviated as MM.
- _____. *Anthropology from a Pragmatic Point of View*, trans. Mary J. Gregor. The Hague: Martinus Nijhoff, 1974. Abbreviated as PA.
- _____. "Religion Within the Boundaries of Mere Reason," *Religion and Rational Theology*, trans. Allen W. Wood and George Di Giovanni. Cambridge: Cambridge University Press, 1996.
- Korsgaard, Christine M. *Creating the Kingdom of Ends*, Cambridge University Press, 1996
- Mahon, James Edwin. "Kant on Lies, Candour and Reticence," *Kantian Review*, 7 (2003): 102-113.
- O'Neill, Onora. *Constructions of Reason*. Cambridge: Cambridge University Press, 1990.
- Paton, H. J. "An Alleged Right to Lie: A Problem in Kantian Ethics," *Kant-Studien* 45 (1953–54), pp. 190–203.
- Potter, Nelson. "Duties to Oneself in Kant's Ethics," *Kant's Metaphysics of Morals: Interpretive Essays*, Mark Timmons (ed.). Oxford: Oxford University Press, 2002.
- Raz, Joseph. *The Morality of Freedom*. New York, Oxford University Press, 1986.
- Schapiro, Tamar. "Kantian Rigorism and Mitigating Circumstances," *Ethics* 117 (October 2006), pp. 32-57.
- Sedgwick, Sally. "On Lying and the Role of Content in Kant's Ethics," *Kant-Studien* 82 (1991). pp. 42–62.

Wilson, Donald. "Middle Theory, Inner Freedom, & Moral Health," *History of Philosophy Quarterly*, 24 (2007). pp. 393-413.

_____. "Moral Health, Moral Prosperity, & Universalization in Kant's Ethics," *Teorema*, (2005).

Wood, Allen W. *Kant's Ethical Thought*. New York, Cambridge University Press, 1999.

_____. *Kantian Ethics*. New York, Cambridge University Press, 2008.

Endnotes

- ¹ See, for example, Christine Korsgaard, *Kant's Formula of Universal Law and Kant's Formula of Humanity* (Korsgaard, 1996); Barbara Herman, *Moral Deliberation and the Derivation of Duties* (Herman, 1996); & Onora O'Neill, *Consistency in Action* (O'Neill, 1990). There are, of course, exceptions, notably Allen Wood who argues that this model of interpretation tends to distort and underestimate Kant's view (Wood, 1999, 2008).
- ² Using Korsgaard's view as an example, Schapiro stresses that Korsgaard thinks of deception "as an act of interference with or usurpation of another's rightful authority to govern himself" that is intended to undermine the victim's attempts to choose for himself how to act. Viewed in this way, deception is regarded as an act that is "necessarily designed to be a way of (illegitimately) interfering with another's autonomy" and as something that is "wrong in itself or by its very nature" (Schapiro, p. 38).
- ³ This rigorism is famously illustrated in Kant's remarks about lying to prospective murderers and has been widely discussed in the literature. See, for example, Mary J. Gregor, *The Laws of Freedom*; Barbara Herman, "Moral Deliberation and the Derivation of Duties" and "Obligation and Performance" (Herman, 1996); Thomas E. Hill Jr., "Making Exceptions without Abandoning the Principle: Or How a Kantian Should Think about Terrorism," (Hill, 1992), & "A Kantian Perspective on Moral Rules" (Hill, 2000); Christine Korsgaard, "Two Arguments against Lying" & "The Right to Lie: Kant on Dealing with Evil" (Korsgaard, 1996); H. J. Paton, "An Alleged Right to Lie: A Problem in Kantian Ethics"; Sally Sedgwick, "On Lying and the Role of Content in Kant's Ethics"; & Allen Wood, "Lies" (Wood, 2008).
- This rigorism is not, however, limited to extreme cases like the murderer example that are usually the focus of discussion in the literature and a comprehensive response here requires a unified account capable of dealing with the nuances of the broader range of problem cases we confront here.
- ⁴ References to Kant's works cite the volume and page number of Kant's *Gesammelte Schriften* (published by the *Preussische Akademie der Wissenschaften* (Berlin: Walter de Gruyter, 1902-). Translations and abbreviations used in the paper are cited in the reference section.
- ⁵ There seems, for example, to be no reason why Kant would think of the moral wrong in self-murder as categorically different from other-murder – if killing a rational agent in one's own person in order to reduce discomfort is morally wrong, killing another person for similar reasons will be wrong also, and be wrong for the same general reason(s).
- This is not to claim that duties like these will necessarily have the same scope. I mean to suggest only that in cases where the CI prohibits the same kind of actions (or similar maxims) involving ourselves or others, we ought to expect these duties to share a common concern, and thus that the wrongs involved ought to be related. (I take this point to represent a natural extension of Barbara Herman's claim that accounts of the Categorical Imperative ought to be didactic, Herman, 1996, Ch. 7)
- ⁶ I have argued elsewhere (Wilson, 2005) that there are reasons to use this work to inform key ambiguous notions in Kant's moral philosophy and to offer a more practical version of his theory. This paper is intended as a continuation of this work.
- ⁷ Stephen Engstrom, "The Inner Freedom of Virtue" (Engstrom, 289-315).
- ⁸ That is, to the extent that we are (negatively) free from causal determination by sensible impulses and positively free in that we are subject to the moral law, we already have the basic capacity to choose to constrain ourselves on the basis of reason characteristic of inner freedom and this notion does not therefore introduce any new dimension of freedom.

⁹ This model of Kant's views on self-deception is widely shared by commentators. Nelson Potter, for example, suggests that, "Self-deception poisons the well of inner awareness (for example, awareness of our own motives of action) and thus undermines in a fundamental way our ability to be accurately aware of our inner self, its motives, and actions" (Potter, 376). In a similar vein, Allen Wood writes:

Those who falsify what they know even before their own minds are undermining their rational functioning in the most basic possible way. When people do it in order to avoid painful consciousness of their own faults and failings or to put their own actions in a more flattering light than they deserve, this also amounts to a subversion of moral conscience at its very foundations (Wood 2008, 257).

¹⁰ As I understand him, Kant therefore treats some lies to others as violations of a duty owed to ourselves because he thinks that some lies like this express or facilitate self-deception. Other commentators interpret Kant's treatment of the duty prohibiting lying owed to *ourselves* differently and think of it as encompassing a broader class of lies.

Allen Wood, for example, argues that the only untruths that Kant thinks of as lies violating duties owed to *others* are lies involving "intentionally untruthful declarations" meant to deprive someone of something that is theirs by right (for example, "a piece of property, or a choice it is their right to make" (Wood 2008, p. 242). According to Wood, Kant treats *all* other lies to other people as violations of the duty prohibiting lying owed to *ourselves*, including, presumably, otherwise deliberate and self-serving falsehoods, manipulative misdirection and omission, etc. that fall short of false declarations infringing on rights.

While space precludes any serious discussion of this alternative, I think that the texts here can also be interpreted in the manner I suggest and that there are independent reasons to prefer a broader understanding of the duty prohibiting lying owed to others that is more in line with the broader scope associated with the deception-as-interference model.

Kant stresses the limited scope of *juridical* duties prohibiting lying owed to others in the Doctrine of Right, emphasizing that an active sensitivity to how others may choose to understand what we say cannot be required of us *in terms of Right* and that the only kind of lie prohibited by *juridical* duty is therefore one that "directly infringes upon another's right (e.g., the false allegation that a contract has been concluded with someone made in order to deprive him of what is his" (MM, 6/238)). In the Doctrine of Virtue, however, he reminds us that duties of right are *also* ethical duties and that when considered in this way these duties are "wide" (in comparison to "narrow" duties of right) and require us to make the right of humanity our end in action (6:390-1). This suggests that as *ethical* obligations the scope of duties prohibiting lying owed to others may be broader, requiring us, for example, to comply with the spirit rather than just the letter of agreements and to be concerned more broadly with other ways in which our actions and discourse can influence and subvert the wills of others. In addition, while Kant's framing assertion that "[T]he greatest violation of a human being's duty to himself regarded merely as a moral being (the humanity in his own person) is the contrary of truthfulness, lying" (6:429) might initially suggest Wood's broader understanding of the duty prohibiting lying owed to ourselves, Wood himself treats this very broad assertion as rhetorical exaggeration and emphasizes that Kant goes on to focus exclusively on "inner lies" and to discuss the duty here narrowly only in terms of self-deception.

I think that considerations like these suggest that there is room for the more limited interpretation of Kant's claim that some lies to others are violations of the duty prohibiting lying to ourselves that I propose (and, correspondingly, for the kind of broader understanding of duties prohibiting lying owed to others that I will suggest below). Questions of interpretation aside, however, I also think that there are independent reasons to prefer this kind of account. Depending on how the idea of false declarations in violation of peoples' rights is understood, the kind of interpretation Wood advocates seems in danger of distorting the moral situation we confront in the case of many lies. We can control peoples' perception of their circumstances in a variety of ways that will be difficult to classify as involving "intentionally untruthful declarations" and we can manipulate their choices of ends in ways that will be difficult to construe as undermining their *external* freedom to act for their own purposes. If so, then actions like these would not constitute violations of duties of non-deception owed to others on Wood's interpretation and

would be reclassified, instead, as violations of duties owed to ourselves. It seems, however, very counterintuitive to treat deliberate and cynical attempts to manipulate others as wrongs done to oneself rather than to them.

- ¹¹ For a more comprehensive discussion of this analysis of the duty prohibiting self-deception see Wilson, 2007. I argue here that the various duties of virtue that Kant discusses in the *Metaphysics of Morals* can be regarded as systematically oriented around concerns with ensuring our capacity for moral agency and that this suggests an attractive and subtle account encompassing a range of psychological and social concerns.
- ¹² I can adversely affect the deliberative capacities of others by drugging them, encouraging and fostering their fears and weaknesses, etc. and I can make rational self-constraint easier for others by doing what I can to ensure a secure public culture conducive to reflection and introspection, and generally by making it easier for individuals to develop and sustain the kind of personal choice of values and commitments in which their inner freedom is realized and strengthened. I cannot, however, make them good. As Kant puts it when describing the different necessary ends associated with duties of virtue owed to ourselves and others:
- [T]he *perfection* of another man, as a person, consists just in this; that he *himself* is able to set his end in accordance with his own concepts of duty; and it is self-contradictory to require that I do (make it my duty to do) something that only the other himself can do (MM, 6:386)
- ¹³ As I interpret him (Wilson, 2007), this is the impetus behind Kant's concerns with vices like defamation, ridicule and arrogance. Kant thinks that the victim of arrogance, defamation or ridicule is likely either to internalize the criticism of others at the cost of her own proper self-esteem or to succumb to bitterness and anger towards others for leveling it. More generally, actions like defamation and ridicule "cast a shadow of worthlessness over our race itself," "dulling" our natural moral disposition by repeatedly exposing us to human frailty, and thus contribute to the creation of a public culture of alienation and contempt incompatible with the respect for humanity "on which the impulse to the morally good rests" (MM, 6:466). Conversely, showing others respect by "softening" our judgments (being tactful and considerate in the way in which errors of reasoning or judgment are brought to another's attention) and sometimes keeping these judgments to ourselves "can arouse their striving to deserve it" (MM, 6:466) and thus improve their condition.
- ¹⁴ Raz, 369.
- ¹⁵ I cannot pursue this here but our obligations to others in these contexts will also be limited by related concerns with our own agency and we will not therefore be reduced to having to spend our days providing information needed by others.
- ¹⁶ Herman argues that the Categorical Imperative is properly applied to generic maxims of self-interest (maxims of the form "I will do x to get what I want") and that when a generic maxim fails this procedure the outcome is a deliberative presumption indicating that the actions involved are not permissible means to ends of self-interest. Deliberative presumptions derived in this way are regarded as informing and governing everyday moral judgment and as setting the terms of occasional moral deliberation (by serving to flag the salient features of situations and the moral considerations they involve). Herman, 1996, Ch. 7.
- ¹⁷ The wrong apparent in this latter case will be viewed here as similar in character and not one regarded merely as a failure of beneficence. Indeed, not only would these both be of concern on the same grounds but if the important matter is one bearing closely on the agent's sense of herself or her organizing commitments there could be reason here to think that the latter failing is more serious.
- ¹⁸ In this respect at least, the standard Kantian model therefore tends to better reflect our normal intuitions about these cases than the consequentialist view of paternalistic interference. Consequentialist arguments tend to emphasize the difficulty of accurately judging another's good and suggest that the consequences will typically be

better if we leave people to make their own judgments about what is best for them. As many critics of the consequentialist analysis emphasize, however, this argument justifies only a defeasible rule of thumb prohibiting paternalistic deception: one that can always be easily overridden in cases where we can reasonably judge to know better what would be good for someone else and that, as such, is not capable of supporting the kind of principle of veracity that commentators like Sissela Bok think underlies the common intuition that truthfulness should be a more robust norm in our dealings with others.

- ¹⁹ Korsgaard, for example, emphasizes the impossibility of clearly distinguishing cases where autonomy has been lost and argues that any dividing lines we draw here will be conventional and ought to reflect facts about our nature and the limits of our understanding and the non-ideal circumstances of our moral lives (Korsgaard, p.357). This might allow us to draw these lines in ways that would encompass some of the problem cases we confront involving adults making bad choices in adverse circumstances but still only by treating the individuals concerned as though they had lost the capacity for autonomous decision making.
- ²⁰ For the same reasons, the inner freedom account will also not protect us from the ordinary adverse affects of others' actions. These will routinely interfere with our ends and activities in ways that impede us and learning how to manage obstacles and inconveniences like these arising externally from the permissible activities of others and the general circumstances of our agency is an equally important part of learning how to manage one's own affairs and function properly as an agent.
- ²¹ This case is adapted from adapted from an example of Thomas Hill Jr.'s (Hill, 1991).
- ²² It is certainly possible that physical and psychological abuse could reduce someone to a condition of abject compliance to such an extent that we might deem them to have lost the capacity for autonomy. I assume, however, that there are also cases of serious abuse that stop short of this extreme in which it would be a distortion to treat the individuals concerned as though they were children without the capacity to decide for themselves. I mean the present case to be understood in this way.
- ²³ It is worth noting here that we would also need to be reasonably confident that any deception on our part will not be thwarted by discovery and that it will otherwise have the desired effect(s). The reason for this is not, however, that the permissibility or justification of my action depends on the likelihood of good consequences coming from my deception. A sincere effort at living up to *any* of our moral commitments requires us to think about the consequences of our actions and try to bring these into line with what is expected of us: If I am serious about my duty to repay a debt, for example, I need to think about, and take, the requisite practical steps to accomplish this goal and not, say, simply drop the money on the lender's doorstep. The point here is simply that if we are to think of any act of deception as being consistent with the principled requirement of respect for others' inner freedom, it needs to be the case that we can reasonably judge that, given the people involved and the way the world normally works, what we propose to do is likely to actually turn out to be compatible with the specific requirements this imposes.