

Self-Constitution: Action, Identity, and Integrity

Lecture Six

Integrity and Interaction

Christine M. Korsgaard

6.1.1 In my last lecture, I described Plato's Constitutional Model of the soul. According to the Constitutional Model, as we saw, reason must rule in the soul, if the soul is to be capable of action, and it must rule according to its own principle, if the action is to be a good one, and not defective. If something other than reason supplies the soul with its ruling principle, then the soul is not ruled for the good of the whole, and this is revealed in the fact that the constitution of such a soul contains the seeds of its own fragmentation. Situations can arise in which the person's will is turned against itself, cannot rule consistently with itself, and so undercuts its own efficacy. The timocratic person may lose track of his ends in his efforts to maintain his honor. The egoistic person prefers an apparent satisfaction to the very reality needed to make sense of that satisfaction. The democratic person drops his projects in the face of the slightest temptation or distraction. The tyrant, as we saw, is different, for he is unified; but being ruled by some end that appears to him to be worth doing *anything* for, he cannot strictly speaking be said to choose *actions* at all. That is, he doesn't decide what's worth doing for the sake of what, because it's already settled that he's going to do whatever is necessary for the sake of this one end. So although he is unified, it isn't in the way necessary for action. But what then *is* the principle that unifies the soul the way necessary for action? What is reason's own principle?

6.1.2 But here we find Plato to be suddenly unhelpful. For what he tells us is that reason's own principle *just is* the principle of unifying the soul. He describes the good person's deliberations this way. (This quotation is rather long so I've put it on the handout:)

One who is just does not allow any part of himself to do the work of another part or allow the various classes within him to meddle with each other. He regulates well what is really his own and rules himself. He puts himself in order, is his own friend, and harmonizes the three parts of himself like three limiting notes in a musical scale - high, low, and middle. He binds together those parts and any others there may be in between, and from having been many things he becomes entirely one, moderate and harmonious. Only then does he act. And when he does anything, whether acquiring wealth, taking care of his body, engaging in politics, or in private contracts - in all of these, he believes that the action is just and fine that preserves this inner harmony and helps achieve it, and calls it so, and regards as wisdom the knowledge that oversees such actions. And he believes that the action that destroys this harmony is unjust, and calls it so, and regards the belief that oversees it as ignorance. (443c-444)

It is as if we asked: which principle could this be, that enables the just person to unify himself? And it is as if Plato replied: don't look for some *further* principle that has unity as an *effect*. Justice is not some other or further principle that enables us to maintain our unity as agents. It is that very principle itself - the principle of maintaining our own unity.

Now consider a comparison. Kant argues that we must will in accordance with a universal law in order to be autonomous. One might ask Kant: which law? Which law

could this be, that enables a person to be genuinely autonomous? And Kant might reply: don't look for some *further* law that has that as an *effect*. The categorical imperative is not some other or further law that enables us to be autonomous. It is that very law itself, the law of giving laws to ourselves.

6.1.3 Now these may be the right answers, but Plato and Kant have not made it obvious, for at this point we reach a famous crux, that shows up in both of the arguments. Kant tells us that action requires autonomy and autonomy requires universalizability. Plato tells us that action requires unity and unity requires procedural or constitutional justice in the soul. Both of them then tell us that these properties are only fully realized in the soul of the morally good person. But how do we get from the formal properties that these philosophers have identified as essential to action – from universalizability and from procedural or constitutional justice – to a substantive commitment to morality?

In Kant's argument, the crux concerns the question whether an action that is autonomous in the formal or procedural sense - in the sense that the agent chooses its maxim for himself - must also be in accord with the categorical imperative in the more substantive sense necessary for morality: a universal law that governs all rational beings, yielding reasons that all of us can share. Why, after all, if the agent is autonomous, should there be any limitation on which maxim he can choose?

And in Plato's argument, the parallel crux comes at a famous or perhaps infamous moment in Book IV. In the text immediately preceding the long passage I just read out, Socrates proposes that he and Glaucon should test his theory by appealing to what he calls the "ordinary cases." Accordingly, he asked Glaucon whether the just person as they have by now described him, the unified person, would embezzle deposits, rob temples, steal,

betray his friends or his city, violate his oaths or his other agreements, commit adultery, be disrespectful to his parents or neglect the gods, to all of which Glaucon says, with a complaisance that is startling to the reader, “no, he would not, the just person as we have described him would not do these kinds of things.” Why, we find ourselves wondering, is the usually skeptical Glaucon so sure?

6.1.4 What is the relationship between maintaining unity in your soul, and doing things like telling the truth, keeping your promises, and respecting people’s rights? Here the Constitutional Model again suggests a way to approach the question. When we call a political state “just” or “unjust” there are in fact two different things we might mean. We might mean that the way the government treats the citizens - and so the way the citizens treat each other - is just or unjust. Or we might mean that the state as a whole is just or unjust in its dealings with outsiders - with foreigners, for instance, and, especially, with other states. I will call these things *inward* and *outward* justice. A state or a city-state is *inwardly* unjust if it is racist or sexist or lacks a free press or has an established religion or has a large population of homeless citizens whom it makes no effort to help - things like that. It is *outwardly* unjust if it is imperialistic or violates treaties or has unfair trade laws or secretly meddles in the affairs of other states.

By the analogy then, Socrates has been describing what inward justice is in a person. When he comes to the test of the ordinary cases, though, he asking whether the person he has described would be outwardly just. It’s interesting to notice that Socrates proposes this as a *test*, that is, he uses it to *confirm* that his description of inward justice is correct. This shows that Socrates takes the connection between inward and outward

justice for granted - he assumes that a person who is inwardly just will be outwardly so. But why should this be so obvious?

6.2.1 Well, here's a case for you to think about for starters. Suppose you are a utilitarian. When other people ask you questions, your answers are guided, not - or at least not in the first instance - by what you think is the truth, but by what you think it is useful for those other people to believe. Of course you think that *in general* it is useful for people to believe the truth, so this doesn't have to mean that you go around telling lies right and left. Nevertheless, you don't tell people the truth until you have satisfied yourself that it is useful for them to believe the truth, so when you answer their questions the first question that you must ask *yourself* is what it is useful for them to believe. And there is your outward justice. And then one day it occurs to you that there is really no reason for you to treat yourself any differently than anybody else. Believing something, committing yourself to a certain way of conceiving the world, is after all, a kind of action, and it is just as true of you that your beliefs and actions should be useful as it is of anybody else. Or try it this way: you are a person, and you don't tell the truth to people until you've satisfied yourself that its useful for them to believe the truth, so you shouldn't tell *yourself* the truth until you've satisfied yourself that its useful for you to believe the truth. So you undertake to believe, not what is true - at least not in the first instance - but what it is useful for you to believe.

But now you've got a problem. For before you can allow yourself to believe what is true, you have to satisfy yourself that it is useful for you to believe what is true. But you *can't* do that first without first satisfying yourself that *what* you think about this - about what it is useful for you to believe - is, quite simply, true. That is, when you say to

yourself “well in this case, it is useful for me to believe the truth,” or as it may be “in this case, it would really be better if I believed something false” you have to be satisfied that those thoughts, the ones I just mentioned, are, quite simply, true. So you have to try to tell yourself the truth, in the first instance: there is no way around it.

And this is no accident. Thought requires a little more respect for the humanity in your own person than utilitarianism allows. For treating yourself in accordance with the Kantian prohibitions against coercion and deception is a condition of thought, without which thought isn't possible. You can't treat yourself in accordance with the principle of utility while you are thinking. For the principle of utility is a tyranny, while thought, by its very nature, is free.

6.3.1 Let's return to the argument. Do inward and outward justice necessarily go together? Following the rules of Plato's analogy, we might pursue this question by asking whether inward justice and outward justice go together in the case of actual political states. Is a nation's honesty and fairness in its dealings with other nations the natural outward expression or consequence of its honesty and fairness in its dealings with its own citizens? Certainly we can spot, in a general way, pressures and tendencies in this direction, although they may not seem decisive. To start with, there are the simple embarrassments of elementary inconsistency - making a big fuss about human rights at home while trampling on them overseas, say - but there other sorts of pressure as well. It's hard to have a free press and yet lie to the world. So is it hard to have a free *mind* and yet lie to the world? This is a question I shall come back to. Here's another case. Let's suppose, as I think we may suppose, that constitutional democracies are inwardly just, and that war is at least usually outwardly unjust. Now Kant argued that constitutional democracies would

not be very likely to go to war, since in a constitutional democracy the citizens have to give their formal consent to a war, and that isn't likely to happen, since it is usually only the rulers, not the citizens, who really want to go war. Constitutional democracies, as we have learned since then, do sometimes go to war, but it's one of everyone's favorite observations these days that they don't go to war *with each other*.

6.3.2 And this brings me back to Plato. It's a curious fact that he has little to say about his perfect Republic's practice of outward justice. It's more curious still that the little he does have to say is contained in one of the most sinister moments in a book that, after all, has no shortage of sinister moments. In the passages I have in mind, Adiemantus is wondering how the Socrates's ideal Republic will be able to fight wars if, as Socrates has claimed, it is not going to be a wealthy city. Socrates proposes that when the Republic is attacked, it should make an alliance against its enemy with some third city, offering the third city the proceeds of victory as an incentive. That's bad enough, but the passage I have in mind is the one that follows, when Adiemantus complains that if the wealth of all the other cities came to be gathered in a single city, the Republic would be in danger. And Socrates replies:

You're happily innocent if you think that anything other than the kind of city we are founding deserves to be called *a city*.

Adiemantus asks what Socrates means, and Socrates says:

We'll have to find a greater title for the others, since each of them is a great many cities, not *a city*... At any rate, each of them consists of two cities at war with one another, that of the poor and that of the rich.... If you approach them as one city, you'll be making a big mistake. But if you

approach them as many and offer to give to the one city the money, power, and indeed the very inhabitants of the other, you'll always find many allies and few enemies (422a-423a).

Socrates here recommends exploiting your enemy's fundamental weakness, turning its factions against each other, winning the allegiance of the poor by offering them the possessions, and perhaps even the persons, of the rich. Consider what the analogy on the personal level would be to *that*. Appeal to your enemy's worst desires, the ones that will expose the disunity in his soul, foment civil war within him, and you can render him incapable of effective action. Is *this* how the good person is going to deal with his enemies? The essential Kantian commitment to respecting another's humanity, even when that other acts badly, seems a long way away. Of course we are in the territory of non-ideal theory here, for we are talking about war, and even Kant agrees that in the case of war the rules are different. But it's still a little chilling.

6.3.3 Yet Plato has a point. How do you interact with someone who is seriously divided against himself? *If you approach them as one city*, Plato says, *you'll be making a big mistake*. To see the problem, consider one of the stock characters of contemporary moral philosophy, a character who is, in fact, a sort of diachronic version of the disunited city - I mean Derek Parfit's nineteenth century Russian nobleman. The story goes like this. The nineteenth century Russian is now, in his youth, a socialist, and he plans to distribute large portions of his inheritance, later, when he comes into it, to the peasants. But he also anticipates that his attitudes will become more conservative as he grows older, and that he may not think that this is the right thing to do, when the inheritance is finally his own. So he makes a contract *now*, to distribute the land when he gets it, which can only be

revoked with the consent of his wife, and he asks his wife to promise not to revoke it *then*, even if he tells her *then* that he has changed his mind, and that she is released from the promise. Parfit makes it clear that the case is not like that of Ulysses binding himself to the mast to resist the Sirens' song. The young Russian does not anticipate that he is going to become irrational, that his judgment will be clouded, or that the immediate temptation of having the estates will undermine his self-control. He simply believes that when he is older he is going to have different values than the ones he has now. Parfit portrays him as telling his wife that his younger self is his real self, that his ideals are essential to him, and that if he loses those ideals she should regard him as effectively dead. Being dead, he cannot release her, and if his middle aged avatar claims to release her from the promise, in an effort to keep hold of the estates, she should regard him as someone else, who therefore cannot release her - almost, although Parfit does not put it this way - as a kind of impostor, posing as the continuation of his younger self. If she does make the promise, and it all happens as her husband predicts - when the estates come to him he wants to keep them after all, and he tries to release her from the promise - then, Parfit says, "It might seem to her as if she has obligations to two different people."

To two different people, it appears, at least one of whom she now *must* wrong. Parfit envisions her deciding to keep her commitment to her husband's younger self, feeling that he is the one she ought to be loyal to, the one that she loves. But staring that fact in the face is an everyday moral reality: she is a married woman, yet now she must set herself up as her husband's enemy, accuse him of being a kind of impostor, and deny him a voice in the disposition of the estates - and she must do all of this, although he is not, as Parfit has specified the example, in any way irrational or out of control. With what possible right could a wife do this? Don't be blinded by the fact that, substantively

speaking, you are probably on the young man's side, and hope that the peasants will get the estates, for there are procedural issues at stake here.

6.3.4 The procedural issues I have in mind spring from the nature of marriage. What is marriage, anyway? This will sound like a digression but you'll see in a moment why I am raising the question. Kant thought that marriage was the solution to a problem. In both the *Lectures on Ethics* and the *Metaphysics of Morals* Kant claims that there is something morally troublesome, even potentially degrading, about sexual relations. What bothers him is *not* the conventional idea that one is using another person as a means to his own pleasure. That's not what Kant thinks, and, in any case, any difficulty about that would, according to Kant's own theory, be alleviated by the other's act of free consent: a person is not using you as a *mere* means if he has your own free consent. What worries Kant is rather that sexual desire takes a *person* for its object. He says: "They themselves, and not their work and services, are its Objects of enjoyment." (LE 162) Regarding someone as a sexual object is not like regarding him as an instrument or a tool, but something more like regarding him as an aesthetic object, something to enjoy. Only in this case the attitude is not just appreciation but desire (MMV 426/87). So viewed through the eyes of sexual desire another person is seen as something wantable and therefore, inevitably, possessable. To yield to *that* desire, to the extent it is really *that* desire you yield to, is to allow yourself to be possessed – in a way, to be someone's property. The problem is how you can do that in a way that is consistent with respect for your own humanity. And the solution, according to Kant, rests in a complete reciprocity of possession, in which each of you belongs wholly to the other. Kant says:

If, then, one yields one's person, body and soul, for good and ill in every respect, so that the other has complete rights over it, and if the other does not similarly yield himself in return and does not extend in return the same rights and privileges, [then] the arrangement is one-sided. But if I yield myself completely to another and obtain the person of the other in return, [then] I win myself back; I have given myself up as the property of another, but in turn I take that other as my property, and so win myself back again in winning the person whose property I have become. In this way the two persons become *a unity of will*. (LE 167; my emphasis)

It actually doesn't matter for my purposes what you think about Kant's view of the problem, so long as you are prepared to believe that he's on to something when he offers the solution: that marriage creates a unity of will. Two people who get married pledge to share one another's lives and possessions; there is therefore a range of decisions that they pledge to make together – decisions about where they will live, decisions about the house, the car, the children, and the money - including the disposition of those inherited estates. Indeed, Kant makes the joint ownership of property one of the necessary conditions of marriage, believing as he does that without equal ownership the absolute equality of power that is necessary for reciprocal possession cannot possibly be maintained. So the Russian nobleman's wife cannot operate as an independent person free to choose now between two loyalties. She has unified her will with that of her husband, and therefore she is committed to making the decision together with him. But how can she do that, when he cannot make the decision together with himself?

6.3.5 And anyway, think about the assumptions behind the young nobleman's original request. The young nobleman asks his wife to commit herself, to make a promise, and to keep her promise in the future. She is to hold him, by holding herself, to giving up the estates. But if she can do this, why can't he? Nietzsche wrote that in order to have "the right to make promises" one must "be able to stand security for her own future" and that this involves having "a real *memory of the will*." For, as Nietzsche points out – I'm quoting now:

between the original "I will," "I shall do this" and the actual discharge of the will, its act, a world of strange new things, [new] circumstances, even [new] acts of will may be interposed without breaking this long chain of will.

Nietzsche believed that the achievement of this kind of will, a specifically human achievement, required an enormous overcoming of our distractible, forgetful, animal nature, and that centuries of human cultural evolution, political development, punishment, and with it suffering – and especially suffering - went into the breeding of– I'm quoting:

the sovereign individual, the person who has her own independent protracted will and the right to make promises -- and in her a proud consciousness, quivering in every muscle, of what has at length been achieved and become flesh in her, a consciousness of her own power and freedom, a sensation of humanity come to completion.

I've been taking liberties with the translation of the pronouns again, for the Russian nobleman's wife is apparently a creature of this kind, or at least the Russian nobleman relies on the thought that she is, but it is all too evident that no such proud consciousness is quivering in *him*. His only way of keeping his commitment is to get his wife to keep it

for him. So what is she supposed to think of his *marriage vows*? Nietzsche recommends that creatures like her – and I quote - “reserve a kick for the feeble windbags who promise without the right to do so.” I don’t know about *that*, but I think we can agree with Plato here: *if she approaches her husband as one person, she’ll be making a big mistake.*

6.3.6 We were looking for the conclusion that someone who is inwardly just will also be outwardly just. Instead, we wandered off into another but related point. People who are not inwardly just are hard to *treat* justly. That’s one of the deepest problems of non-ideal theory. Inward justice in the people involved is a necessary condition of outward justice in human interaction.

6.4.1 What is human interaction, anyway? Kant’s theory of marriage is just one instance of a general theory of human interaction. Friendship, as Kant understands it, is also the solution to a problem. The problem is how you can devote yourself to the happiness of someone else without losing track of your own. And as in the case of marriage, the solution involves a reciprocal exchange that leads to a unification of the two friends’ wills. Kant characterizes friendship in the *Lectures on Ethics* as “the maximum reciprocity of love.”(LE 202) There he argues that friends *exchange* their private projects of pursuing their own happiness, each undertaking to care for the other’s happiness instead of his own. Kant says: “I, from generosity, look after his happiness and he similarly looks after mine; I do not throw away my happiness, but [I] surrender it [into] his keeping, and he in turn surrenders into my hands” (LE 203). This requires the maximum reciprocity of love because, as Kant says, “if I am to love him as I love myself I must be sure that he will love me as he loves himself, in which case he restores to me that with which I part and I come

back to myself again” (LE 202). So the structure is just like that of marriage. I pledge myself to pursue my friend’s happiness, but her happiness in turn includes my own; she pledges herself to pursue my happiness, but mine now includes hers. So like marriage, the exchange produces something new, a shared object, *our* happiness, which we now pursue together, and make decisions about together, as the object of our unified wills.

6.4.2 But it is not just these ongoing relationships that according to Kant involve the formation of unified wills. It is everyday interaction itself. For the idea also shows up in Kant’s treatment of two closely related issues in the theory of political right. Kant conceives all rights on the model of property rights, and that means that he supposes that when we make someone a promise what we are doing is in effect giving that person a kind of property right in an act of our own. To make a promise is therefore is to transfer your right over your “choice to perform a specific deed,” as Kant puts it, to another. Your action, or at least your choice whether to perform the action, now belongs to the other and the other can claim it. So if I promise you that I will meet you for lunch tomorrow, a certain act of my will – the decision whether to meet you for lunch – now belongs to you and not to me. That is why it is you and only you who can release me from my promise – because the choice now in fact belongs to you.

But there is a problem, as Kant points out, about how this can possibly happen. Here’s the problem. Suppose I offer you a promise. Until you accept my promise, I can always take it back - I am not committed until you have accepted. But if in making my offer I have not yet committed myself, then I have not yet promised, and you have nothing to accept. If promises were empirical, this problem would be insoluble, for there is necessarily a temporal gap between my offer and your acceptance, and however small

the gap is it prevents the promise from happening. The same problem exists and is even more vivid when what is happening between us is a transfer of an ordinary piece of property rather than of an act of my will. How do I transfer my property to you? It cannot be that I first abandon my property and then you pick it up, for if I abandoned it then during the interval anyone could legitimately pick it up - it would be unowned property, which anyone may claim. Yet until I have abandoned it is mine, so how can it possibly become yours? If you take it when it's mine you'll be stealing it. The temporal gap again prevents the transfer. These problems show that promises and transfers cannot be understood as the result of successive acts. Instead they must involve from a single common will, from a moment of unity between us. Promises and transfers involve four acts, Kant says: the empirical offer and the acceptance, which are actually mere preliminaries, and then the two reciprocal acts of will that constitute the unification of our two wills that makes the transfer possible. You agree to take what I agree to relinquish, I agree to relinquish what you agree to take, and so we form a unified will that the thing in question should be yours and not mine: it all has to happen at once, as a single action, if it is to happen at all. Kant even says that we try to symbolize this unification of our two wills by performing essentially simultaneous empirical acts such as *shaking hands*.

6.4.3 What does it mean to say that we unify our wills? If you think of the will as a pre-existing entity, rather than as a product of self-constitution, it sounds very mysterious, like some sort of act of ontological fusion. But Kant doesn't mean anything like that. When we interact with each other what we do is deliberate *together*, to arrive at a shared decision. Since the conclusion of a practical syllogism is an action, the result is an action that we perform together, governed by a law we freely choose together. The free choice of this law

is an act that constitutes our unified will and makes shared action possible. That, in Kant's view, is what interaction is.

6.5.1 So the possibility of interaction depends on the possibility of shared deliberation. And that possibility in turn depends on a certain conception of reasons. Our reasons must be what I will call public reasons, reasons whose normative force can extend across the boundaries between persons. Public reasons are roughly the same as what are sometimes called objective, or agent-neutral reasons. They may be contrasted to what I will call private reasons – subjective or agent-relative reasons. A private reason is a reason whose normative force is private, in the sense it belongs only to one person. As many philosophers have pointed out, privacy is consistent with a kind of universalizability requirement. On the private conception of reasons, if I think that the fact that something will make me happy is a good reason for *me* to do it, then universalizability requires me to think that the fact that something will make you happy is a good reason for *you* to do it. But my happiness is still mine, and yours is still yours; mine is a source of reasons for me, but not for you; yours is a source of reasons for you, but not for me. A public reason, by contrast, is a reason that is shared.

6.5.2 If interaction in Kant's sense is possible, I must treat the other person's reasons, as I will put it, *as reasons*, that is, as considerations that have normative standing, normative force, for *me*, and therefore as public. And to the extent that I must do that, I must also treat the other as an end in himself, that is, as a source of reasons, as one whose will can be legislative for me. To see why, consider a simple coordination problem. Suppose you and I are related as teacher and student, and we are trying to schedule an appointment. "Stop

by my office right after class,” I say, thinking that that will be convenient for me, and hoping that it will also be convenient for you. It isn’t, as it turns out. “I can’t,” you say, “I have another class right away.” So I have to make another proposal. It’s important to see why I do have to do this: it’s because having the meeting is something that we are going to do together. The time I suggested isn’t good for *you*, and therefore it isn’t good for *us*, and it follows from that that it isn’t after all good for *me*, and so I need to suggest another time. To perform a shared action, each of us has to adopt the other’s reasons as her own that is, as normative considerations with a bearing on her own case. That’s why the fact that the time is not good for you means that it also is not good for me. So we both keep making suggestions and considering them until we find a time that’s good for both of us. The aim of the shared deliberation, the deliberation about when to meet, is to find (or construct) a shared good, the object of our unified will, which we then pursue by a shared action. And it follows from the fact that the action is shared that if either of us fails to show up, we will both have failed to do what we set out to do.

6.5.3 Now let us consider the alternative. Suppose I don’t treat your reasons as reasons, as considerations with normative force for me. What does that mean? We might think it means that the discussion will go like this: I ask you to meet me right after class. As before, that isn’t possible. “I can’t,” you say, “I have another class.” Perhaps I just look puzzled as if I cannot make out why this is supposed to be relevant and I say something like “well, skip it.” I don’t try to work out something that’s good for both of us, because your reasons are *nothing* to me, and any objections that might be grounded in your reasons are nothing to me: I can’t see why your reasons should count; all that counts is

that a certain time is good for me. Obviously, we can't relate at all on those terms. So if that's how it is, no interaction is going to be possible.

6.5.4 But of course that is putting it too strongly, for I was imagining that I disregard your reasons altogether, as if they weren't even there. But there is something in between treating someone else's reasons as reasons, and disregarding them altogether. I could take your reasons into account, not as public reasons with normative implications for me, but as private ones with normative implications for you, implications that bear on my predictions of how you are going to act. I know that you have certain reasons, and that being rational you are likely to act in accord with them; this fact is one of the facts I have to deal with in planning my own actions. I don't in this case treat your reasons as reasons, instead I treat them as possible tools and obstacles, things that might help me to achieve my ends or get in my way. As I said before, if I treat your reasons as reasons, they may change my mind about what counts as the best outcome. Only an appointment time that suits us both is the best outcome, and that is what I am aiming at, since that is what our shared deliberation is aiming at. But if I treat your reasons as tools and obstacles, they will come into my deliberation in a different way – I will see whether I might use them as tools in the pursuit of my own ends, or if not, if they are obstacles, then I will try to determine whether I can remove them from my path.

In my example, your reason not to meet me right after class presents itself to me as an obstacle. So I might try to get rid of it: I might try to persuade you to skip your class, say to talk you out of thinking you really have the reason that is in my way, or that it has as much force as you thought. "Is something *important* happening in your class that day?" I might ask. Or perhaps I will hope I can just mow your reasons down, by bullying.

“Look,” I say, “there’s really no other time we can meet.” If these techniques don’t work, I may change my mind about what is the best possible outcome. But it won’t be like the case in which I treat your reason as a reason, and find that a certain outcome isn’t best for me because it isn’t best for us. If I change my mind about when to meet because your reasons present me with an obstacle, it will be because given the obstacles that stand in my path, finding a new time is the best I can do. So if I regard your reasons this way, as tools and obstacles, we will be able to have a *sort* of relationship – for we will be able to fence and negotiate and bargain. Well, isn’t that how it is? Certainly it must be admitted that when you *are* trying to arrange a meeting with someone who cannot meet at any time that is convenient for you this is often *exactly* what it feels like.

But it is one thing to say that it is *work* to overcome the egocentric predicament, and another to say that we can’t do it. For we must keep in mind it is also *work* to overcome the temptation of the moment, and yet sometimes we must, if action is to be possible at all. Is interaction really just a matter of negotiation and fencing and bargaining between two private reasoners? Consider that if reasons are really private then there is no reason for the sort of negotiation I have been describing to be open and above board. Quite the contrary: insofar as I am a rational private reasoner, I should use force or tricks if I can, because I see your reasons only as obstacles to be defeated, or tools to be used. In effect, this means that our relationship will be a kind of war, or combat. We aren’t aiming at constructing shared reasons, so one of us, the one whose private reasons will prevail, will be the victor, and the other will be the vanquished. And in trying to achieve that victory we will rightly use all the arts of war. Or of course we may have to reach a compromise, but we won’t think of the compromise as a good thing in itself, because it is

something that respects both of our reasons as far as that is possible. Instead we will each think of it, privately, as the best he or she could do under the circumstances.

6.5.5 And this of course is how the Russian nobleman is related to himself. He doesn't think of his future reasons as reasons – he thinks of them as facts to contend with, as tools and obstacles, and in his case mainly obstacles – and he is therefore is in a condition of war with himself. His efforts as a young man are dedicated to insuring that his younger self wins, and his older self loses. His soul is therefore characterized by civil war, and that is why he fails as an agent. And for the same reason he is unable to interact with his wife.

6.6.1 Anyway, what does it mean to say that a reason is private, mine, not yours? On a Kantian conception of reasons, the claim will be ambiguous. Every reason arises from the endorsement of a proposal presented by an incentive. When we talk about *my* reasons we may mean the reasons that I endorse, the ones that I legislate, the ones that I embody in my actions. This notion *is* ineluctably first-personal, and there is nothing wrong with that, for the notion of human action, self-conscious action, is of a course a first-personal notion. But the first person doesn't have to be the first person singular: the fact that a reason has to be mine in this way doesn't prevent it from being ours. So taken in that way, the category of *my* reasons doesn't exclude an identification between my reasons and the reasons of others.

But sometimes when we talk about “my” reasons or a person's own reasons, we mean the ones that arise from a particular set of incentives, some set of incentives that is supposed to be in some plain sense mine. The most obvious candidates for this are the incentives that arise from events in my body – the appetites, pains, and pleasures to which

I am necessarily, as an embodied being, subject. But we can add to those, in a vague way, any incentives that a person might experience in the natural course of his individual embodied existence. Now I think it is true that there is a clear sense in which these incentives are a person's own, for her agency and her embodiment are necessarily related – it is only in so far as a person is embodied that she can act at all. When she legislates for her self, she is legislating for – let me put it this way for the moment – all the conscious inhabitants of her body, present and future. And although there is room for controversy here I am prepared to say that as an embodied being she *must* take into account all of the incentives, present and future, to which she will be subject in the natural course of her embodied life: that is part of what Plato means when he talks about willing for the good of the whole. But to say that *only* the incentives that arise directly in me in the course of my individual embodied existence can be the source of reasons for me is simply to beg the question against the possibility of interaction. I *must* interact with the conscious inhabitants of my body, because I must act with my body. But I *may* also interact with other people, and when I do then their reasons, as well as my own, become as it were incentives in the deliberative process that we undertake together, resources for the construction of our shared reasons. So taken in that way, the category of *my* reasons doesn't exclude an identification between my reasons and the reasons of others. I need not legislate alone, and that being so, I need not regard the incentives that arise in the natural course of my individual embodied life as the only possible sources of reasons for me.

6.6.2 Now the defender of private reasons will claim that I have said nothing so far that he needs to deny. For of course the private reasoner will say that he can choose to interact with some other people. He may decide to deliberate together with certain others, for

various private reasons of his own. But until this decision is made he has no reason to accord normative force to the reasons of others. So what are these private reasons for genuine interaction going to look like? Perhaps in view of certain ends that he and another person share he thinks the other would serve well as an ally; or perhaps he and the other have made an agreement to take each others' reasons into account; or perhaps he simply cares about some other people – he loves or respects them - and taking their reasons to be normative for him is part of what it means to care in this way.

It's hard to see how the first of those reasons, sameness of purpose, could launch the private reasoner out of his solitary stance by itself. Suppose we both have a certain purpose. If your purpose coincides with mine, then you and your reasons are usable as tools for me, and I and my reasons are usable as tools for you, and our awareness of this may give us a sort of reason to act together. But it does not really give me a reason to accord your reasons any normative standing of their own, although it may give me a reason to pretend to, in order to keep you on my side. Well, suppose then we move to option two, and we make an agreement? We agree to pursue our shared end together, and we therefore agree that when we deliberate about how to pursue it, we will give each other a vote, so to speak, in the deliberative process. So here a private reason gives us each of us a reason for treating the others reasons as normative considerations, and our reasons, yours and mine, as reasons that are public between us. But there are problems with this proposal. For one thing, if we make the agreement for some specific purpose, and that purpose is achieved, then we seem to run into Hobbesian problems: why shouldn't I free ride on the agreement, for whatever private reason drove me into it in the first place? And why shouldn't you? And if I know that, why should I trust you? and so on and so on... by now we all know how *this* goes. It is the familiar stuff of the crime caper – once the

robbery is successful and the spoils are in our hands then our previous agreements about the division of the spoils is up for grabs again. So long as I treat your reasons as having normative force only because it somehow serve my interests to do so, I am not really treating them as reasons after all, because if my stance towards them is instrumental I am necessarily regarding them as tools. In order to treat your reasons as having authentic normative standing I must treat them as having normative standing simply because they are *your* reasons, out of some sort attitude of love or of respect that I have towards you, and for no other reason than that.

So this brings us to the third option, that you interact with some people because you care about them. But that isn't really a case of deciding to interact with someone for a private reason. It doesn't work that way: it is not that I decide that because I love or respect someone I will accord normative force to his reasons. It's rather that finding myself loving or respecting someone, I simply do accord normative force to his reasons. The interactive stance may be called forth in us by something about a person or about our relationship to a person. But it is not, therefore, something chosen.

But set that point aside, and say that the private reasoner can choose to interact with some people he cares about. Would that show that morality is optional, depending as it does on whether you have any private reasons that favor interaction? It is not that simple, for there is one person with whom even the most determined private reasoner must interact. And that is himself.

6.7.1 Interaction, I have argued, is quite literally acting with others. But for a creature who must constitute her own identity, it is equally true that acting is quite literally interacting with oneself. The requirements for unifying your agency internally are the

same as the requirements for unifying your agency with that of others. Constituting your own agency is a matter of choosing only reasons you can share with yourself. That's why you have to will universally, because the reason you act on now, the law you make for yourself now, must be one you can will to act on again later, one you can live with later, come what may, unless you come to see that there's a good reason to change it. This is what Plato means – and he means it literally – when he says that a good person is his own friend, and legislates for the good of his soul as a whole. The Russian nobleman fails as an agent because he doesn't do this, because he doesn't will a law he thinks can commit himself to acting again on later, come what may. In fact he has what in lecture two I called a particularistic will, for he expects to change his mind without a reason. And this fact is the *same* fact as his lack of self-respect. In saying this, I don't exactly mean to imply that he has to respect his future self, or at least I don't mean that he has to respect his *future* self's verdict about the proper disposition of the estates, although that is certainly one way that he could go. That is, he could decide that his future self, in deciding to keep the estates, is making a claim with normative standing, and that if he and his future self are to act together he must take that claim into account. He could decide that he must not give up the estates, or perhaps not all of them, without a manifest failure to will for the good of his future self, and so for the good of the whole. Ordinary prudence with respect to an overly enthusiastic charitable impulse might look like that. But we don't have to saddle the Russian nobleman, or ourselves, with some dreary piece of ordinary prudence in order to make the case here, for, as I said, his respect for his future self is not what is at stake here. This isn't because his future self has no standing, but rather because his future self is just himself. He can decide to disagree with his own future attitude. But *unless* he is then also prepared to regard his own future attitude as one of weakness or irrationality, he is

not according to the reason he himself proposes to act on *right now* as having normative standing. For he is not making a law for himself unless he thinks of his future attitude as a violation of that law, and if he does not think he can make laws for himself then he lacks self-respect. So his problem is not his disunity with his future self, but his disunity with himself here and now. And his problem is not disrespect for his future self, but disrespect for himself here and now.

6.7.2 I have suggested that the conditions for successful interaction are the joint conditions of respect for the other's humanity, and the treatment of her reasons as considerations with public normative standing: when we interact, we legislate together, and act together, for the good of the whole we in this way create. But action is simply interaction with the self. If this is so, then respect for one's own humanity, and the consequent treatment of one's own reasons as considerations with public normative standing, are the conditions that make unified agency possible. Without respect for your own humanity it is impossible to will the laws of your own causality, to make something of yourself, to be a person; and unless you make something of yourself, unless you constitute yourself as a person, it will be impossible for you to act.

6.7.3 There is one question that this argument, even if it works, may seem to leave open, and I have time here only to gesture at the answer. I have argued that respect for your humanity is a condition of action, because you must interact with yourself, and respect for the humanity of those you interact with is a condition of interaction. I have also argued that authentic interaction with others is possible, though only possible under the condition of respect for their humanity and the treatment of their reasons as public. We

are not locked away from each other in the space of private reasons. In the only senses in which reasons are private, we are not excluded from interacting with others. Nor do we need a reason to interact, as the private reasoner supposes: the strategic stance towards is not the default position. But it may not be clear that you must interact with others, as opposed to taking up an essentially strategic stance towards others – that genuine interaction is the default position. I do take that to follow from the argument, for respect for the humanity in your own person is a condition of action, and that respect will be called forth by the humanity in others as well. But isn't it only *your own* humanity you must respect? This is an ill-formed question. What is your own, in the individual sense of your own, is not your humanity but what you make of it, your practical identity, and the existence of that *depends* on your respect for humanity in general.

If this is right, respect for the humanity in your own person, and that of every other, is a necessary condition of effective action. It enables you to legislate a law under which you can be authentically unified, and it is only to the extent that you are authentically unified that your movements can be attributable to you, rather than to forces working in your or on you, and so can be actions. So moral law is the law of the unified constitution, the law of the person who really can be said to legislate for himself because he is the person who really has a self. It is the law of successful self-constitution.

6.7.4 So the basic insight behind Plato and Kant's confidence here is not really anything surprising or paradoxical. It is simply that every person interacts with others as he interacts with himself, and in this the good person is no different. A person who cannot keep a promise to himself cannot keep a promise to another. A person who is prepared to sell himself for a little money is prepared to sell others as well. A servile person lacks

respect for his own rights and so for the rights of others. Inward and outward justice go together. Earlier I suggested that it is hard for a state to have a free press, and yet lie to the world. So is it hard to have a free mind and yet lie to the world? The answer of course is yes. For when the good person thinks, she tries to tell herself the truth, and when she talks, she is just thinking in the company of others. And more generally, for a creature who has to constitute herself – to make up her mind and her will – acting is simply interacting with the self, just as interacting is simply acting with others. It is this that made Plato and Kant so confident that the inwardly just person will also be outwardly just; or to put it Kant's way, that legislating for oneself, and legislating for the Kingdom of Ends, are one and the same thing.