

Responsibility, Moral and Otherwise
(The Routledge Lecture, 21 February 2013)

Philosophers have been concerned with and puzzled by the nature of responsibility at least since Aristotle's time. What does it take to be a responsible agent? Under what conditions is it justifiable to hold someone responsible for her character or behavior? Notoriously, it seems to many that genuine responsibility requires a kind of metaphysical freedom that is hard to understand much less believe in, while others insist that no such metaphysical freedom is necessary.

A common way of introducing the subject calls attention to two senses of responsibility. Sometimes when we say "X is responsible for Y" we mean only to name a causal connection: X is an individual or a state of affairs that makes a salient causal contribution to the occurrence of Y. In this sense, your cat might be responsible for the spilled milk, the rain might be responsible for spoiling the picnic. By contrast, this paradigmatic introduction continues, when we charge an individual with moral responsibility, we mean something more. Like the cat, a person can be causally responsible for the spilled milk if she knocks over the bottle during an epileptic seizure; like the rain, a person can be causally responsible for spoiling the picnic if in the course of it she is stung by a bee and goes into anaphylactic shock. But a person might also spill the milk on purpose, wanting to ruin her host's tablecloth or disrupt the family brunch; she might spoil the picnic by her irritable carping or by rudely insulting the other guests. In these latter cases the person's connection to the event under discussion appears to reflect more deeply on *her*, and, at least initially, it does not seem inappropriate to blame her. These latter cases

thus seem to exemplify a different and deeper kind of responsibility that persons, and perhaps only persons, may bear.

The familiar contrast between causal responsibility and moral responsibility however, seems to me to blend two apparently independent distinctions together. On the one hand, we may distinguish merely causal responsibility from another kind, which is deeper. On the other hand, we may distinguish our responsibility for the *moral* qualities of our actions and our character from our responsibility for traits and behavior that are of primarily nonmoral interest. The traditional, or, as I earlier called it, notorious philosophical problem of freedom and responsibility seems to me to be concerned with the difficulty of accounting for the depth of the connections we make in charging people with responsibility for their actions and traits. By identifying deep responsibility with moral responsibility, we risk muddying the waters by blurring concerns about depth with debates about the nature and basis of moral demands.

Thus, for example, one increasingly popular view understands moral responsibility as accountability for meeting standards we impose on each other of mutual respect and concern. We expect people to have or display a certain degree and kind of good will to each other, and we blame them if they do not. Since the epileptic who spills milk because of a seizure is not expressing ill will (or anything at all), he is not morally responsible and in particular not blameworthy for the event, whereas the intentionally disruptive milk-spiller does express an attitude that, according to this view, would make it appropriate to hold her morally to account. As

an account of when a person is liable to specifically moral praise or blame, this focus on the good or ill will of the agent is important and insightful, but it sidelines a more general concern about the nature of responsible agency. If there is such a thing as being deeply responsible for nonmoral character traits, actions, or their consequences, this account of moral responsibility will have nothing to say.

What can be said in favor of the idea that deep responsibility includes more than responsibility for the moral qualities of one's actions and one's self? What can be said against it? This is what I want to explore in this paper. My hope is that the inquiry itself may help us clarify what we mean by, or want from, responsibility, and that it may help us make progress on identifying the features that lie at responsibility's core.

It will be best to proceed through a consideration of cases. Roughly three categories of examples come to mind. First, is a range of traits that are either of primarily prudential interest or that fall on the continuum of what some have called executive virtues and vices, traits that help or hamper one's ability to realize whatever other ends or values one has. Laziness and industriousness, cowardice and courage, gluttony and temperance, all fall within this range. A second range of traits worth considering are those nonmoral traits not included in this first category that make a person more or less appealing to himself or others. If you ask yourself what you most like and dislike in yourself, in your friends, or in your acquaintances, I suspect that many of the traits you come up with will belong to this second category. Having or lacking a sense of humor, being charming or pretentious are

good examples. With some ambivalence, I shall call this category of traits ‘aesthetic’, though it is important to remember that such traits need not have anything to do with beauty or the arts. Of course, character traits have consequences. What you say, when you laugh, how you spend your time, will reflect all sorts of nonmoral as well as moral traits. When considering whether we are deeply but nonmorally responsible for anything, we should think not only about whether we are responsible for the *traits* in the two categories I have just mentioned, but also about whether we are responsible for the *behavior* and other phenomena that result from a person’s possession of such traits. The third category that I would like to consider in exploring the idea of nonmoral responsibility is art, broadly understood. Thinking about the extent and sense in which we hold artists responsible for their artistic creations may reveal interesting parallels to and differences from the way we think about people’s responsibility for the moral good and harm they do.

An initial case for the idea that we should understand deep responsibility to extend beyond the moral can be made by noting the parallels between the sorts of considerations that alternatively encourage or discourage us from holding people *morally* responsible for the consequences of their actions and those that make us more or less likely to judge them in nonmoral ways. For example, imagine that you are pushed by a stranger, rudely, as it seems to you, and so you are initially inclined to blame him. If you learn that he was pushed or hypnotized or blackmailed, you withdraw or lessen your blame. Similarly, we may imagine that you walk into the home of an acquaintance and see that it’s a mess – clothes on the floor, dishes in the sink, piles everywhere. Initially, you think “what a slob!” But you withdraw your

judgment and whatever attitude accompanies it, if you learn that his home was ransacked a few hours ago, or that his arthritis makes it particularly hard for him to pick up after himself. Alternatively, you might be attracted or engaged by a painting your friend produced and so be inclined to credit him, until you learn that the painting was a result of his tripping over some cans of paint he had foolishly left open on the floor. The contrast between a merely causal sort of responsibility and a deeper kind seems as clear in the latter cases as in the former.

Reflection on these and similar cases suggests an analysis of the distinction between causal and deep responsibility that can easily accommodate the analysis of moral responsibility that I earlier criticized as too narrow. For, just as the question of whether to hold someone morally as opposed to merely causally responsible for pushing you seems to depend on whether it is right to interpret his act as an expression of his good or ill will toward you, so the question of whether to hold him deeply, as opposed to merely causally responsible for his messy home or his attractive painting, depends on the degree and the way in which these things express his choices and more particularly his values or point of view. Although a person may be *causally* responsible for anything that results from a feature of her being or behavior, she is only *deeply* responsible when the result expresses or reflects what we may be tempted to call her soul, or, to avoid any suggestion of metaphysical or religious commitments, her (inner) self. Gary Watson has called the conception of responsibility that takes one to be responsible for actions insofar as they are governed by and expressive of one's self "the self-disclosure view", and I

will follow him on this.¹ Since one's self presumably includes more than one's attitudes toward others and more than one's moral (and immoral) values, such a conception allows one a sphere of responsibility that stretches well beyond the moral and, at least by contrast to merely causal responsibility, may be considered deep.

The view gains support from consideration of a contrast first articulated by P. F. Strawson in his seminal article "Freedom and Resentment." Wanting to steer our thinking about freedom and responsibility away from its too close association with the justifiability of reward and punishment, Strawson reminds us of the attitudes, like resentment and indignation, gratitude and admiration, that the practices of reward and punishment frequently express. He suggests that we understand responsible agency roughly as the kind of agency that makes one an appropriate target of such "reactive attitudes," by contrast to the kind of agent to whom, we think, only "the objective attitude" is justified.

"To adopt the objective attitude to another human being," Strawson writes, "is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps simply to be avoided, though *this* gerundive is not peculiar to cases of objectivity of attitude. The objective attitude may be emotionally toned in many ways, but not in all ways," he continues, "it may include repulsion or fear, it may

¹ Gary Watson, "Two Faces of Responsibility," *Philosophical Topics* 24, 1996 p. 228.

include pity or even love, though not all kinds of love. But it cannot include the range of reactive feelings and attitudes which belong to involvement or participation with others in inter-personal human relationships; it cannot include resentment, gratitude, forgiveness, anger, or the sort of love which two adults can sometimes be said to feel reciprocally, for each other. If your attitude towards someone is wholly objective, then though you may fight with him, you cannot quarrel with him, and though you may talk to him, even negotiate with him, you cannot reason with him. You can at most pretend to quarrel, or to reason, with him.”²

Strawson himself emphasizes the reactive attitudes of resentment, indignation, and gratitude – attitudes that, as he points out, are reactions to the degree of good or ill will their objects seem to bear toward us, and are therefore particularly tied to the target’s moral character and behavior. For this reason, the accounts of responsibility I criticized as too narrow are often called “Strawsonian accounts.” But the range of reactive attitudes that, as a group, stand in contrast to the objective attitude must be more inclusive. The distinctive kind of interpersonal love that Strawson explicitly includes as among the reactive attitudes, for example, is typically not based exclusively on the degree of good will the beloved has towards us. It may be based also on his sense of humor, his political engagement, his attitudes to nature, his taste in art.

² P.F. Strawson, “Freedom and Resentment,” in Gary Watson, *Free Will* Second Edition (Oxford: Oxford University Press, 2003) p. 79.

As we saw in the long passage I quoted, the reactive attitudes as a group are connected with the ability to see an individual as a *participant*, or as a potential participant in a certain kind of relationship with one, as someone one can reason with, appealing to a sense of 'reason' that goes beyond negotiation. In fact, I suspect that 'reason' is too cold and too narrow a word to fully capture what Strawson is aiming at. But in any case, it seems clear enough that the contrast between those to whom reactive attitudes are appropriate and those toward whom we ought have only the objective attitude will roughly place human adults of normal mental health and intelligence in the first class, while placing most lower animals, human infants, and adults with certain severe mental deficiencies in the second.

I mentioned earlier that the attitudes we might bear to an artist on the basis of his or her artwork display interesting parallels, if also certain contrasts from the attitudes we form toward a person on the basis of his morally good or bad deeds. As my example of the accidental painting showed, we only give a person credit for producing a work of art if he did it on purpose – just as we only feel grateful (or think we ought to feel grateful) to a person who has benefited us if she did so intentionally. Indeed, just as my gratitude to you depends not only on whether you benefited me intentionally, but on whether you benefitted me out of good will, and not, for example, because you were hoping to impress my brother, or to get my vote, the credit we accord an artist depends not only on his producing the painting on purpose but on his producing what is good or interesting about the painting in a way that is not merely accidental. If, as I have heard it rumored, El Greco suffered from a vision problem that made him insensitive to the distortions of the figures he

painted, it compromises, or at least calls into question, the degree to which he deserves praise for the unsettling but soulful qualities of his painting. (Arguably, it may call into question whether his paintings are soulful at all.)

But the comparisons that interest me have less to do with our tendencies to give praise, credit, or their opposites to artists on the basis of their work than with our tendencies to react in other ways – to like or love some artists on the basis of their artwork, though having a chillier response to others.³ Speaking for myself, there are some novelists and painters that I love (qua artist, that is, as opposed to qua potential friends or lovers)⁴ while others, whose work I recognize as equally impressive, leave me cold. I love Henry James, but not James Joyce, Matisse but not Picasso. These attitudes (roughly, of love or chilliness) seem to me as *reactive* as resentment and gratitude, as much in contrast to the objective attitude as indignation or respect. The judgments and feelings I bear toward James on the basis of his novels depend at least as much on my taking them as expressive of his soul, as disclosive of his self, as do the judgments of blame or praise and feelings of resentment or gratitude that I might bear toward someone who has wronged or

³ It is noteworthy that, while it is may be almost automatic to praise, credit and admire an artist for producing good work (assuming that the artist's connection to the quality of her work is not accidental), we are unlikely to blame someone for being a bad artist or for producing inferior art unless we think she has the talent and skill to do something better and so see her work as evidence of her pandering to an audience for the sake of some goal of which we disapprove. In that case, however, our criticism or blame is too close to being moral itself to serve as a possible object of comparison.

⁴ In the sense I am interested in, it is possible to love Caravaggio, qua artist, or even Francis Bacon, without necessarily thinking it would be nice to go out to dinner with him.

benefitted me if I think he has done so responsibly. This suggests that when Strawson draws the contrast between the reactive attitudes and the objective attitude in terms of the former's association with a *participant* stance, we would do best to understand "participation" broadly, as referring not necessarily to membership or even potential membership in a literal common community (all the artists I named are dead, after all), much less to a community whose sole purposes are moral and political. We should rather understand the participant stance as involving the idea that the individual in question is 'one of us' in some wider or other sense – as being, perhaps, psychologically like us, or at the same level, in having the same sort of self or soul to be expressed.

The considerations that I have offered lend support not only to the idea that there is such a thing as deep but nonmoral responsibility; they lend support, in particular to the self-disclosure view of what deep responsibility is. It might give us pause, however, to note that with respect to at least some of the nonmoral traits and behaviors for which we are inclined to hold people deeply responsible in this sense, questions about control, the ability to do otherwise, not to mention metaphysical determinism - questions that have dogged philosophical discussion of moral responsibility since its beginnings - seem forced if not entirely out of place.

Thus, for example, I have said that my feelings, judgments and attitudes toward Henry James reflect the fact that I take him to be deeply, and not merely causally, responsible for his novels. But to ask, in this context, whether James could have done otherwise than write *The Portrait of a Lady* seems utterly bizarre. What

could such a question mean to establish? Is it asking whether he had to write precisely the words he did rather than some others? (Might he, for example, have named his protagonist Catherine instead of Isabel?) Or are we being asked to speculate on whether he might not have written novels at all but gone into manufacturing or dentistry instead? These questions seem irrelevant to the justifiability of the attitudes I form toward James on the basis of his novels. Perhaps, then, we should interpret the question as asking whether, given that James was going to write novels, it was necessary that he write *this sort* of novel rather than one more like *Ulysses*, for example, or *Portnoy's Complaint*. This question seems somehow closer or more relevant to the issue of the justifiability of my attitudes towards James. But the answer I am immediately inclined to give to this question – namely, that it is highly likely that he *did* have to write his sort of novel - seems to me to take nothing away from my tendency to love and admire him in just the way I do.

At least some of the reactions we have toward people based on aesthetic traits of character seem similarly impervious to questions about alternate possibilities and lack of control. Having a good sense of humor, for example, seems to reflect something deeper about a person than the quality of his voice or the color of his hair, and my attitude toward someone – particularly, how much I like him - is considerably affected by his possession of a sense of humor or lack of one. But, before writing this paper, it never occurred to me to ask whether people have control over their senses of humor, and now that it does occur to me, it seems quite

likely that they do not. Recognizing this does not, however, make the possession of a sense of humor matter less to me, nor do I see any reason why it ought to.

Some compatibilists might take these observations as confirmation of what they have thought all along – namely, that the relevance to responsibility of the ability to do otherwise tends to be greatly exaggerated. But others, no less reasonably, will draw an opposite conclusion that calls into question the idea that we are nonmorally but deeply responsible for such things as aesthetic traits and artistic creations after all. The fact that typically we don't blame people for being humorless or boring – and that, at least partly, we don't blame them because we assume that they can't help it – lends support to this more skeptical response, for the fact that we don't blame them seems indicative of our not holding them responsible after all.

One way to account for this mixture of reactions to people's nonmoral traits and their consequences is suggested by Gary Watson in "Two Faces of Responsibility." According to Watson, the self-disclosure conception of responsibility marks one important distinction, between those things that we can and those that we cannot attribute to the person in a certain relatively deep way. But, Watson suggests, something further is required if, in addition to *attributing* a trait or consequence to someone, we also want to hold him *accountable* for it – if we want, for example, to take him to task for it, to expect him to defend himself or apologize or suffer deserved punishment. Because holding someone accountable involves making him liable to blame and punishment, considerations of fairness

come into play that are irrelevant to contexts where only attributability is at issue. In particular, it seems unfair to hold someone liable to blame or punishment for a trait or a consequence that is not under his control.

Thus, it appears that in order for someone to be responsible for a trait or a state of affairs in a way that makes him accountable, the self-disclosure conception of responsibility is inadequate. Fairness requires that we do not hold someone accountable for a trait or state of affairs unless he has control over the thing; unless, in particular he has the ability to avoid having the trait or being associated with the state of affairs for which we are inclined to blame or punish him. This suggests a two-level conception of deep responsibility. At the first level, which is captured by the self-disclosure conception, one is entitled to *attribute* a trait or state of affairs to a person. To say that a person is responsible for something, at this level, is to say that it is disclosive of the person's self. *Accountability* for traits and states of affairs is afforded at the second level, which requires, in addition to self-disclosure, control over that aspect of the self that is relevantly disclosed.

Several features of this two-level conception are worthy of note. First, insofar as it makes sense to think of the first level of responsibility as "deep" in comparison to merely causal responsibility, it seems natural to understand the second level of responsibility as deeper still. For the second level requires not only that the actions or traits whose responsibility is in question issue from or disclose the subject's self, but that the subject be in control of the relevant aspect of her self and the actions or states that issue from it; we might think of this control in terms of

one's self being governed by a still deeper self. Second, and related to this first point, we may note that the debates about the relevance of determinism to freedom and responsibility seem most plausibly to enter in at this second level. This is the level at which control is necessary;⁵ this is the level at which it may be insisted that the subject must have the ability to be or do otherwise. The metaphysical problems involving free will and responsibility concern the proper or relevant interpretation of precisely these conditions.

Finally, we may observe that the two-level view provides a ready explanation and defense of the tendency that I referred to at the beginning of this paper – namely, the tendency to conflate (or, less contentiously, to identify) deep responsibility with moral responsibility. For, according to this view, we need a deeper relation between a subject and a trait or state of affairs in order legitimately to hold her accountable for it than we would need in order merely to attribute the trait or state of affairs to her. But “accountability” might plausibly be thought to be a specifically moral notion, or at least one that only or primarily comes into play with respect to the moral qualities of character and action. Since the deepest kind of responsibility is that which is necessary for accountability, and since it is only in connection with moral successes and failures that the question of accountability arises, there would be nothing lost by identifying deep (or deepest) responsibility with moral responsibility.

⁵ At least, regulative control. (See e.g., Fischer, John Martin and Ravizza, *Responsibility and Control*. (Cambridge: Cambridge University Press, 1998). Thanks to EJ Coffman's for reminding me of their distinction between regulative control and guidance control.

Though the two-level view of responsibility has considerable appeal, I believe it is mistaken. Although I agree that the kind of freedom and responsibility one needs in order legitimately to be held accountable for something is different from what is needed for something to be properly attributed to one's self, I don't believe that the former kind *builds upon* the latter, or that it is necessarily deeper. What depth there might be to our status as responsible beings, a depth we are prone to interpret in metaphysical terms, is an issue that pertains as much or more to the kind of self to whom we want to *attribute* actions and character traits as to the kind of self whom we want to hold *accountable* for the consequences of these actions and traits. Whether we are ever as deeply responsible for anything as our practices and attitudes seem to presume, therefore, is a question, I believe, that can be as fruitfully asked about our nonmoral successes and failures as about our moral ones.

To explain why I think the two-level view of responsibility is mistaken, let us look more closely at a train of thought that at first seems to support that view. It begins by comparing two cases of morally objectionable behavior. In the first case, a young man takes the last or the best piece of pie for himself. He acts selfishly, but, given his youth and the fact that his parents have spoiled him, we judge that he could no more help his selfish character than he can help his sense of humor (or lack of it). In the second case, the greedy pie-taker is older and, we think, ought to have known better. In both cases, we might say the pie-taking is attributable to the pie-taker. It is not, after all, as if either man was brainwashed or coerced or didn't know what he was doing. But, given the way I have described the two cases, we may be

inclined to hold the second man more deeply responsible for his action. He, and perhaps only he, is properly held accountable and blameworthy.

The two-level view can account for this by noting that while both agents are responsible at the first level, only the second man is responsible at the second. This is because the first man cannot help being selfish, whereas the second is understood as having more control.

One difficulty with the train of thought I have just presented is its implicit assumption of an unrealistic relation between a person's character and his acts. For, presumably, the reason we are inclined not to hold the younger man accountable for his pie-taking is that once it is stipulated that he ought not to be blamed for *being* selfish, we conclude that it would be wrong to blame him for *acting* selfishly. We think that if his character is not under his control and his action flows from his character, then his action must not be under his control either. But the very fact that people frequently act *out of* character should be enough to call this conclusion into doubt. The fact that I am honest doesn't guarantee that I literally cannot tell a lie; the fact that I am shy doesn't mean that I can't force myself to mingle at a party. Thus, even if it is true that a person cannot help *being* selfish, it wouldn't follow that he cannot be accountable for a selfish act.

In any case, what seems really to be doing the work in supporting our inclination to blame the second man and not the first is the stipulation that the second man ought to have known better – he ought to have known not to take that piece of pie. In fact, on reflection it really doesn't matter why he actually *does* take

the pie – it doesn't matter whether he takes it out of selfishness or malevolence, or even whether he takes it to bring it home to his wife. What matters is simply that (1) he acts as he ought not to, (2) he was in a position to know that he ought not to so act, and (3) it was up to him whether to act in that way or not. Whether his pie-taking discloses a selfish self, a narcissistic self, a malicious self, or even an overly wife-doting self, doesn't matter to his being accountable and blameworthy for the act at all.

Of course, the *quality* of the blame – the more particular negative attitude we take toward the man – whether we think, for example, “that selfish bastard” or “that insensitive pig” will depend on these details. But the quality of the reactive attitude has nothing to do with whether he is accountable. The kind of responsibility that is to be identified with self-disclosure comes apart from the kind that is concerned with accountability.

This point can be made in an even stronger way by considering the case of a psychopath, or at least of a possibly fictitious and inaccurate conception of what some philosophers and psychologists have taken a psychopath to be. I have in mind a kind of person who lacks the capacity for empathy, or for any sympathetic feeling for another creature's misery or pain, who therefore has no conscience, but who is nonetheless able to control his behavior in accordance with the dictates of prudence or any other type of value. Such a person, I imagine, cannot help being a psychopath (or a sociopath, as this sort of character is also sometimes called). More than that, if it is right to think of psychopathy as a kind of disorder, rather than as a character

trait or a vice, then it seems to me that we ought not to hold reactive attitudes toward psychopaths (at least not on the basis of behavior that is informed and shaped by their psychopathy). One may see them as dangers, even as enemies, but only in the sense of “enemy” that might equally apply to man-eating tigers or sharks. It seems to me, then, that a psychopath is not only not accountable for being a psychopath – his psychopathy is not even attributable to him.

But it is not at all clear that this implies that holding him accountable for his acts is out of the question. If he acts criminally, for example, knowing that the acts are forbidden, and that he will be punished for them if he gets caught, and if, as I have stipulated, he can control his behavior, then it does not seem unfair to impose the penalty: he took a chance, and he got caught. In this case, even more clearly than in the earlier one, the kind of responsibility necessary for accountability comes apart from the kind necessary for attributability. For, the psychopath, according to this suggestion, may be accountable for his psychopathic behavior without his psychopathic character being attributable to him at all.

I suspect that the relative independence of the two kinds of responsibility is masked by the fact that “blame” seems capable of referring to both. When you blame someone you hold him accountable, subject to penalty, to the demand for compensation or apology, and so on, but, typically, you also have a reactive attitude toward him based on what you take to be the character he has disclosed. What needs to be questioned is the assumption implicit in the two-level view that the conditions for the first kind of responsibility for the action depend on and go further

than the conditions of the second kind of responsibility for what the action expresses.

Once we separate the two kinds of responsibility, rather than see the one as building upon the other, what becomes of the metaphor of depth that seemed a natural gloss on the kind of responsibility referred to as moral responsibility that philosophers were so used to introducing in contrast to merely causal responsibility? Does it remain appropriate to think of the responsibility that is necessary for accountability as requiring something deep? I have no doubt that some people will think so, on the grounds that it is unfair to hold someone accountable – and in particular, negatively accountable– for something that it is metaphysically impossible for him to avoid. I am not so sure. For if liability to punishment, the demand for compensation or apology, and so on, are separated from the licensing of reactive attitudes, and if the kinds of punishment and other demands at issue are restricted to what seem to be the only kinds that we humans (as opposed to a divine ruler) have the authority to exercise anyway, then it is not clear to me that anything beyond the ordinary garden varieties of avoidability are required. Thinking again of the criminal psychopath, it seemed that in order for him to legitimately be held accountable, all that was needed was that he knew what penalties would accompany what actions, and was able to choose what to do in light of them. It is not obvious that the ability to choose just referred to requires anything metaphysically robust.

The two-level view arose because the self-disclosure view offered a conception of responsibility that didn't seem deep enough to capture all of our intuitions about what full moral responsibility involved, including those intuitions that have suggested to many that full responsibility required metaphysical indeterminism. But, once we replace the two-level view with a more straightforwardly disjunctive view, the second kind of responsibility – that which is necessary for accountability, may no longer seem especially deep - at least, so my discussion of the psychopath would suggest. Was the idea that there was something deep about moral, as opposed to causal, responsibility just an illusion then?

I believe that if we turn back to the self-disclosure conception of responsibility we will see that it is more complicated than we initially thought, and that if we consider it further, the idea that the kind of responsibility it is meant to capture is genuinely deep may reappear. For though we have said that a person is at least somewhat more deeply responsible for something if it can be attributed to him as disclosing his self, we have so far said very little about what a self is, that will explain why certain features of a person are to be regarded as part of his or her self, while others are to be understood as outside of it. Nor have we said anything that explains or marks out what it takes for an individual to have (or to be) a self of the sort that can be regarded as responsible in this first sense at all. Do nonhuman animals have selves, for example? Do children? For that matter, do psychopaths?

I earlier suggested that a person's self includes her values and her point of view, but this is inadequate as a characterization of the self. For one thing, it doesn't tell us what features of a person are *excluded* from her self. For another, further reflection on the case of the psychopath suggests that even the idea that one's values and point of view are always to be included in the self may need to be revised. For although the psychopath, as I am imagining him, is constitutionally unable to empathize, this doesn't deprive him of the ability to have values or the possession of a point of view. Without revision of the self-disclosure view, therefore, the view seems to imply that insofar as a psychopath's actions disclose his values and point of view, those actions are attributable to him and we are entitled to form reactive attitudes toward him on their basis. But I also suggested that we ought *not* to have reactive attitudes towards a psychopath (at least not attitudes based on actions that are significantly formed or shaped by the subject's psychopathy). If you agree with this, you will think that the self-disclosure view needs to be revised.

How *should* we define or characterize the kind of self that can be a fit object of reactive attitudes? How should we determine what particular features of a person or of her actions are fit bases for particular reactive attitudes? It is too late in the day, and, to be frank, also too early in my thinking about these questions for me to propose, much less defend, a clear and precise answer to these questions. But I do have a hunch, and I will conclude my paper by sharing it with you.

My hunch, to put it briefly, is that the kind of self that is a fit object of reactive attitudes is an "intelligent" self - a self that can sense, understand, and appreciate

the world in the same way (and as well as) we can, a self, in other words, that has the same (or better) powers of perception, reason, and imagination that we do, so that when she responds to the world, it is the same world as ours to which she responds.

This fits with and explains why, as it seems to me, we should take the objective attitude towards the psychopath. Lacking the capacity for empathy, the psychopath is unable to fully understand what other people are like, and so anger, indignation or resentment at his callous indifference to us seems to be misplaced. The psychopath 'doesn't get it' – he *can't* get it, and so we shouldn't be dismayed that he has no concern for us. We should protect ourselves, as we would protect ourselves from sharks, but we shouldn't feel hurt or insulted.

This hunch similarly explains why our judgment of El Greco should change if we learned that astigmatism made him insensitive to the elongated style of his portraits. Eye condition or not, El Greco would still have produced those haunting paintings, and produced them intentionally. But if what he saw in those paintings was not and could not be what *we* see in them, it seems a mistake to react to him, to judge him on the basis of what we see. Moreover, it seems to be a mistake whether our reaction and judgment would be positive or negative.

This hunch also coheres with Strawson's association of aptness for reactive attitudes with a participant stance. For the requirement that the relevant kind of self must have the same kinds of powers of perception and understanding that we

do is a way of insisting that they share - that is, participate in - the same subjective as well as the same objective world as ours.

If one were to interpret the self-disclosure view of responsibility in the way that my hunch would suggest, would the resultant view be able to make sense of our tendency to think of responsibility as deep? Would it be able to capture the association between responsibility and freedom that has made generations of philosophers wonder whether it was compatible with determinism?

Though I am uncertain, I think there is a chance that the answers to these questions is yes. For the abilities to approach and respond to the world with reason and understanding are at the core of my conception of responsible agency, and we frequently refer to the successful use of these powers with words like “profundity” and “depth.” Further, a case can be made that when we exercise these powers in our interactions with the world, we behave more *freely* than we do when our beliefs and values result from prejudice or habit. Finally, it seems to me plausible that one reason determinism seems threatening to freedom and responsibility is that it is difficult to see how determinism can be reconciled with the exercise of genuine faculties of reason, perception, and imagination. And so it may be that there is a kind of responsibility, different and deeper than causal responsibility, that consists in the manifestation or disclosing of an intelligent self’s unimpeded response to the world.

I opened this lecture with the suggestion that the contrast philosophers commonly make between causal and moral responsibility may conflate two

different distinctions, between superficial and deep responsibility on the one hand and between moral and nonmoral responsibility on the other. I further suggested that it might prove fruitful for understanding the concept of responsibility, to turn our attention to the range of nonmoral traits and actions for which we might be deeply but nonmorally responsible. This lecture has mainly been an attempt to convince you of the value of that enterprise. Whether or not it will yield a useful perspective on the concept of responsibility perhaps remains to be seen. But even if it does not, it may provide a helpful and more accurate perspective on ourselves.

Philosophers of action and of ethics tend to think that moral responsibility is *a* central if not *the* central feature of human beings that distinguish us, in a good way, from lower animals and machines. But if moral responsibility is not a part of some larger or more general feature of human agency, it will be irrelevant to our capacity for humor or creativity or to our susceptibility to nature or to beauty. It will be irrelevant to much of what makes us alternatively lovable or obnoxious to each other. This suggests that either moral responsibility has more limited significance than these philosophers think, or – as I would prefer – that the most important and deep kind of responsibility that distinguishes us as human is not limited to the moral.

Susan Wolf
University of North Carolina
Chapel Hill, North Carolina