



This is a repository copy of *Intrusive Intervention and Opacity Respect*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/106299/>

Version: Accepted Version

Book Section:

Bennett, C.D. (2018) *Intrusive Intervention and Opacity Respect*. In: Birks, D. and Douglas, T., (eds.) *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice. Engaging Philosophy*. Oxford University Press, Oxford. ISBN 9780198758617

Intrusive Intervention and Opacity Respect, Bennett, C.D., *Treatment for Crime: Philosophical Essays on Neurointerventions in Criminal Justice* edited by David Birks and Thomas Douglas, 2018, reproduced by permission of Oxford University Press
<https://global.oup.com/academic/product/treatment-for-crime-9780198758617?cc=gb&lang=en#>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Intrusive Intervention and Opacity Respect

Christopher Bennett

ABSTRACT

I claim that coercive neuro-interventions on offenders violate requirements of respect. I explore this idea with reference to Ian Carter's notion of 'opacity respect' – that is, the idea that a form of opacity that Carter calls 'evaluative abstinence' is a necessary feature of respect. I argue that opacity is a necessary part of relating to one another as equals. This is not to say that we should pursue ignorance, or even pretend it. As I interpret it, it is rather the claim that the respect structurally necessary to some inherently valuable form of human relations sometimes requires that we not acknowledge what we see, or could see. The demands of that way of relating to one another are therefore violated when we inquire into the interventions that would be necessary to alter someone's behaviour for the better, and when we deploy the knowledge thus gained in subjecting them to a programme of behaviour-modification.

**KEYWORDS: RESPECT; NEUROINTERVENTIONS; CRIME; OFFENDERS;
REHABILITATION**

1. Introduction

A common theme amongst Romantic thinkers was that the modern pursuit of knowledge can be destructive. 'Our meddling intellect,' Wordsworth famously tells us, 'Mis-shapes the beauteous forms of things:--/We murder to dissect.'¹ These Romantics recommend a form of epistemic abstinence if we are to keep hold of what is beautiful, mysterious, or sublime in our world. There are some things, in other words, that it is better for us not to inquire into.

We can distinguish two ways of interpreting this theme. On the one hand we might read it as a conservative suspicion of modern ways of thinking, coming from a perspective on which tradition, 'naturalness,' and perhaps the idea of social order as a form of the order of nature, are valued for their own sake. This interpretation can lead us in the direction of irrationalism and mysticism. But it would be wrong to think that this exhausts the resources of this line of thought: Romanticism has a more progressive, even revolutionary, legacy. Thus the

¹ 'The Tables Turned,' originally published in his *Lyrical Ballads* (1798).

second way of interpreting this theme is, not as a protest against science and rationalism as such, but rather against a too-simple form of rationalism that applies the same model of knowledge to all areas of human and non-human life. In particular, the view I am interested in is that certain forms of human relationship that are valuable in their own right become unintelligible, undermined, and eroded in the face of social changes of which modern forms of rationality and inquiry are a prominent part. The appropriate attitude, on this second interpretation, is certainly not to reject modernity wholesale, since it brings with it previously undreamed-of opportunities; and certainly not to reject what modern science tells us about the world and our own nature. Rather what we need to do is to recognise the proper limitations of the use of such knowledge, and if necessary, to organise collectively to protect what is valuable from erosion by the forces modernity has unleashed. This can require, amongst other things, a kind of epistemic abstinence. Specifically, it can require us to keep different forms of inquiry to their appropriate spheres, rather than assuming without question that one model will answer to all our complex needs and aspirations.

This paper seeks to apply something like this second interpretation to the question of whether serious offenders should be subjected to crime-preventing neuro-interventions. I don't go so far as to argue that the imposition of such interventions could never be, all things considered, what must be done. In this paper I don't engage with the question of whether there is ever a threshold above which the consequences of failing to subject some offender to moral enhancement would be too awful. Rather my interest is in the picture of human relations in which prohibitions against such interventions have their life. I aim to point out the moral cost that normally serves to prevent such an option from being put on the table, and would be incurred even in such an urgent situation as to make this kind of intervention necessary. What the progressive interpretation of the Romantic theme of 'epistemic abstinence' suggests is that sometimes the respect structurally necessary to some inherently valuable form of human relations requires that we not acknowledge what we see, or could see. The demands of that way of relating to one another are therefore violated when we

inquire into the interventions that would be necessary to alter someone's behaviour for the better, and when we deploy the knowledge thus gained in subjecting them to a programme of behaviour-modification.

In this paper, I explore this idea with reference to Ian Carter's notion of opacity respect – that is, the idea that opacity is a necessary feature of respect.² In the following section (section 2) I set out the target of our discussion. In section 3, I introduce Carter's claims about the relation between respect and epistemic abstinence. In section 4, I develop Carter's view in some key respects, and in section 5, I apply it to neuro-interventions for criminal offenders. In section 6, I consider a range of objections, focusing on the concern that Carter's view proves too much. Section 7 concludes. Two initial qualifications are briefly worth noting. Firstly, the aim of the paper is not definitively to establish the case for opacity respect, but to stimulate discussion by pointing to a neglected way of articulating the case against neuro-interventions for criminal offenders. Secondly, although this is a suggestive theme that, if valid, presumably generalises to various realms of human relations (and state-citizen relations), in this paper I restrict my interest to criminal justice.

2. Crime-preventing neuro-interventions: what; why; and on whom?

In this section, we look at what crime-preventing neuro-interventions might involve. Questions arise about what, who and why: what kinds of intervention would be involved; who would be targeted for such intervention; and what the justification for doing so would be (and whether the justification is sufficient).³

The question of what such intervention involves is not my major focus in this paper, but it might be useful to note some important issues that would affect the

² I. Carter, 'Respect and the Basis of Equality,' *Ethics* 121 (2011), pp. 538-571.

³ In addition, many people will think that there is a further question to be answered: by what right? This last question arises for those who think that it does not simply follow from the fact that it would be good for the state (or any powerful actor) to do something that it therefore has the right or permission to do such a thing: for a recent example, see C. H. Wellman, 'The Rights-Forfeiture Theory of Punishment,' *Ethics* 122 (2012), pp. 371-393. This paper argues that limits to the state's right to intervene arise from the nature of the good that the state aims to realise for citizens: the good of a collective way of life in which individuals participate independently as equals.

question of justification. It would be important to know, for instance, whether it would involve some one-off irreversible change, or whether it is a temporary measure that must be administered repeatedly (and if the latter, what are the conditions that might be placed on its repeated use – for instance, can a person qualify, through good behaviour, to put an end to its use?). Is the intervention surgical or pharmaceutical? What aspects of our psychology would be targeted by such intervention, and how fine-grained can the targeting be? What are the risks of malfunction or side-effects of the intervention? Is it voluntary or compulsory? If voluntary, are there any costs – such as longer term of imprisonment, refusal of parole, post-imprisonment sanctions such as employment restrictions, etc – imposed on those who refuse? For the purposes of this paper, and in order not to make uncharitable assumptions about my opponent’s position, I will assume a model on which interventions are pharmaceutical, and where the effect of the drug is not one-off but must rather be administered repeatedly. This allows for use to be reviewed and curtailed if the subject’s circumstances change sufficiently. However, I will assume that the use of the drug is compulsory. It is not simply offered (non-coercively) to the subject as a form of assistance, but is rather prescribed as part of the sentence.⁴

The question of who should receive such interventions is also not my main focus, but again raises some problematic issues that I note briefly. I assume that some sorts of coerced neuro-intervention are accepted as legitimate when dealing with those who have been diagnosed as insane. Although such treatment can also be controversial, the main question I am dealing with in this paper is whether interventions that are common in the case of those deemed to have serious and crippling mental health problems should be deployed against individuals who would not be similarly exempt from criminal responsibility.

If we ask who might be a legitimate target for such treatment, perhaps the most obvious target category is ‘convicted offenders.’ However, criminologists working on ‘desistance’ tell us that most offenders have a rather shallow,

⁴ This raises a further question: who decides on the use of this intervention in a particular case? A judge? Medical personnel? Elected politicians? Juries?

temporary engagement with criminality that ends, around their early-to-mid twenties, as they start to form families and take on more demanding responsibilities.⁵ The bulk of crime, so the story goes, is committed by a small core group of repeat offenders.⁶ Therefore perhaps it is this latter group that should be targeted, if any are. On the other hand, even in this group there will be plenty of people who commit crimes that, though distressing for their victims, may be relatively minor. So perhaps targets for enhancement would be those convicted repeat offenders whose crimes are serious in the sense of causing serious harm. However, once the element of harmfulness has been introduced, it might seem that it is a person's tendency to cause harm that is the crucial point in making them a suitable target for intervention, and make us wonder whether criminality as such is a distraction. There are perhaps two ways of understanding a criminal conviction for the purposes of legitimating neuro-intervention: on the one hand as evidence of personality defect; or on the other as the normative process through which the person convicted forfeits or otherwise loses a right to non-interference that would normally exclude such intervention. If we take the former interpretation, we might worry that not all convicted serious offenders are necessarily dangerous. After all, there are mistaken convictions, even of supposedly repeat offenders; furthermore, the question can be raised whether all criminal actions are morally unjustifiable, and whether, even were all the crimes on the books prima facie serious wrongs, the machinery of legislation and courts is apt to capture the nuances of justification, excuse and ultimate culpability. Furthermore, it is obvious that there are people who knowingly cause great harm without ever appearing in a criminal court, or suffering conviction. Some of these people have great lawyers; some of them do very harmful things – like gambling on futures – that are not currently illegal; some of them engage in what is labelled 'tortious' or merely 'anti-social' behaviour that stops short of criminality. If we want to target harmful personality, criminal conviction will be a highly imperfect measure. However, the

⁵ J. H. Laub and R. J. Sampson, 'Understanding Desistance From Crime' *Crime & justice* 28 (2001), pp. 1-58

⁶ See for instance Home Office, *Confident Communities in a Secure Britain: The Home Office Strategic Plan 2004-08* (London: Home Office, 2004).

prospect of systematically reviewing the population at large for their propensity to cause harm, and subjecting them to compulsory intervention regardless of conviction, is presumably not a welcome one.

The target position in favour of neuro-intervention that we will assume for the purposes of this paper, then, is that there is some way in which criminal conviction involves the offender losing basic rights to non-interference that he previously had, and that this makes it permissible to assess and intervene on convicted offenders in a way that it is not permissible with the general population. To be maximally charitable, we will assume that it is only those convicted of serious repeat offences that are potential targets of neuro-intervention. Further, to avoid the problems associated with mistaken convictions (where continually protesting one's innocence may look like lack of repentance, and hence greater callousness) and other imperfections of the criminal process, we will also assume that it is only those serious repeat offenders who are independently assessed as dangerous by psychological experts.

Turning to the why question now, the answer to this may be fairly simple. Neuro-interventions seek to make people more moral, or at least to make them such that their external behaviour conforms better to moral standards, in order to reduce harm to (innocent?) people who would otherwise become victims. The reasons the state might do this are either because it takes it as an end in its own right that it should help its citizens towards moral perfection; or it might rather view its interventions in regard to the moral progress of its citizens as a means to its ends of security and justice (that is, the protection of fundamental rights and interests). Since many people think that a liberal state should only have the latter, less ambitious, aims, we will assume this in what follows. The crucial point to make at this stage is that this is an end that only extreme libertarians would reject as illegitimate, and which the state pursues in various projects aimed at promoting good citizenship, including but not limited to public education programmes (about health, the environment, voting and so on). At any rate it will be no part of the argument of this paper that this is an illegitimate state

purpose. So the point of subjecting some to neuro-interventions is to protect potential victims who would otherwise be seriously harmed or killed.

The argument, then, will turn on whether compulsory neuro-intervention is a permissible means by which the state may pursue its legitimate ends of security and justice. My argument will be that neuro-interventions can be contrary to principles of respect – and, crucially, maintaining respectful distance – adherence to which is constitutive of the relationships amongst citizens and the state having a particular kind of inherent value. The next section outlines the view of respect recently developed by Ian Carter.

Before we get on to this point, let me prepare the ground by pointing out what my argument is not. My argument is not that such interventions are problematic because they are coercive, or because they violate absolute rights to bodily integrity, or even to mental integrity; it is rather because of a prior step that explains why (at least some instances of) coercion, or invasions of bodily or mental integrity, are wrong in those cases in which they are wrong. Such interventions are wrong, I will argue, because and insofar as they would have to be guided by knowledge that it is impermissible for us to use as a basis for action; although we may possess such knowledge, and although it might be permissible to undertake to gather such knowledge, principles of equal respect rule out acting on it. This, I believe, may be a more promising explanation of what is wrong with neuro-interventions than simply pointing to their coercive or intrusive character. For although coercion, for example, is often wrong, it is permitted in many circumstances, for instance when the purpose of the agent doing the coercing is important enough. Similarly, although physical integrity is important, it is a moral barrier only up to a point: if you were carrying an infectious and deadly disease that you would spread to the population if not stopped, but which would be easily and harmlessly treated if you would submit to it, and yet you refuse, the argument that ‘my body is my body’ does not, I am happy to accept, hold much weight; the state can and should intervene to stop great harm to others.

What of one's right to one's own mind, or to mental integrity?⁷ The issues here are less clear, partly because it is unclear, when there is so much external causal influence on one's mind independently of neuro-interventions, what a right to mental processes that are authentically one's own would amount to. I'm inclined to think that there is something more to be said here, something which would rely on the connection between rationality and authenticity, and which could therefore explain, for instance, why pharmaceutical interventions are different from other processes by which one's beliefs can be altered, such as persuasion and argument – or even serendipitous events that present new opportunities out of the blue – since the latter can all be thought of in principle as aiding one in one's quest to make sense of one's life, whereas the former – perhaps – cannot.⁸ Thus it may look as though neuro-intervention does, whereas the ordinary causal basis of thought does not, make it impossible for me to think of my thought processes, decisions, etc., as my own spontaneous responses to reasons, or my own sensitivity to reasons. And hence for that reason neuro-interventions undermine our integrity as agents searching to make sense of their lives. Nevertheless this 'mental integrity' line of argument would have to deal with the familiar concern that many types of stimulants and depressants, such as coffee, alcohol and tea, are taken by many not to pose any serious threat to one's search to make sense of things; these things can strengthen that search rather than constituting an abandonment of it.⁹ This suggests that those who want to argue for the value of mental integrity will have to find something to say about the appropriate means one can deploy in trying to make sense of one's life, where one is constrained to use means that are compatible with the terms of the quest; and use the distinction thus drawn to explain the difference between drinking tea and submitting to neuro-intervention. (But then, why should one search to make sense of one's life from where one starts, where one's starting point may be a pretty bad one in comparison to others – can't we use neuro-interventions

⁷ For a recent exploration of these issues, see C. Bublitz, 'Moral Enhancement and Mental Freedom,' *Journal of Applied Philosophy* 33 (2016), 88-106.

⁸ On the nature and importance of this 'quest,' see e.g. A. MacIntyre, *After Virtue* (London: Duckworth, 1981), Ch. 15. We don't have to accept everything in MacIntyre's outlook, of course, to think that there is something suggestive in this idea.

⁹ D. Wasserman and M. Liao, 'Issues in the Pharmacological Induction of Emotions,' *Journal of Applied Philosophy* 25 (2008), pp. 178-192.

to help equalise people's chances of making sense of their life?) Any appeal to authenticity will have to give an account of how we are to identify the authentic self that lies there waiting to be discovered.

At any rate, my line of argument in this paper is different. Whereas the mental integrity argument would take a first-personal perspective, and argue that there is some loss to the agent in taking the apparent short-cut to enhancement that pharmaceuticals or neuro-interventions represent, the present argument takes a second-personal perspective. It claims that certain epistemic constraints are constitutive of a certain mode of relating to another person; when we abandon those constraints, we no longer relate in that way, and something inherently valuable between us is lost. Where the parties so relating are an individual and the state, and where the state, as in a democracy, represents the people as a whole, this undermines the integrity of relations among citizens. The polity can no longer claim to be founded on mutual respect.

3. Carter on equality and opacity respect

Let us now outline the relevant notion of respect, and its connection to epistemic abstinence. I follow Ian Carter's recent work in setting out this notion of respect. Carter's main concern is with the basis of equality, and although, in contrast to his, our approach assumes rather than attempting to establish the importance of equality, it will be easiest to follow his exposition.

The view that 'we' are moral equals is a key modern political value; Carter asks what it is in virtue of which we are equal. The problem he addresses can be put in the form of a dilemma. Any attempt to find the basis of equality in some particular descriptive property would need to find some property that we all possess equally. But this is unlikely because empirically instantiated properties are scalar and hence will be possessed variably by individuals in the group. If moral treatment were to track the grounding property, it would not give us equal treatment as a result. Faced with this problem one might reject this horn of the dilemma, denying that we need to find a grounding property that we all possess equally. After all, we might say, the equality in question is moral equality, not

descriptive equality, so why assume that there is a single property equally possessed by all those we treat as moral equals? But then this impales us on the second horn, since it now looks as though we have given up the search for a qualifying basis for moral equality altogether: we are no longer able to identify the non-moral properties upon which moral equality supervenes.

Carter follows Rawls in arguing that the right way to respond to this dilemma is to locate the basis for equality in a range property. A range property is not scalar but binary – either you possess it or you do not. One possesses a range property when one has some scalar property within a specified range (e.g. rational capacities above a certain level). Thus although we are each capable of rationality to different degrees, any given person ought to be treated as a moral equal if she is rational *enough*. Of course, where we place the threshold is crucial, and one would need a further theory to answer that question. But we are only interested in the shape of the account at the moment.

Carter argues that there are two problems facing Rawls's account. First of all, we need an answer to the question of why the range property (rather than the scalar one) is a relevant qualification for treatment as a moral equal – that is, why it (rationality enough) rather than its basis (rationality) is the key thing. Second, given that rationality is important enough to ground equality, variations in rationality look as though they will still be relevant to questions of how someone should be treated, morally speaking. Thus the variability of the morally important grounding property looks as though it will be a potential challenge to the levelled-out range property when deciding about how to treat someone: one might be faced with a decision about whether respect demands variable or equal treatment even if the demand for respect supervenes on the same empirical base.

The important move in Carter's account is to see that the solution to both of these problems lies in arguing that respect involves accepting that one has reason not to take variations above the threshold as reasons for action. To amend the Rawlsian view, Carter therefore puts forward the idea that treating

someone with respect involves 'evaluative abstinence – that is, a refusal to evaluate persons' varying capacities.'¹⁰

'The kind of respect I have in mind involves ... adopting a perspective that remains external to the person, and in this sense holding back from evaluating *any* of the variable capacities on which her moral personality supervenes, be they capacities for rational thought or capacities for evaluative judgement or capacities for awareness and understanding of one's place in the world.'¹¹

Carter's proposal echoes the 'progressive' strand in Romanticism with which we started. Relating to someone as an equal is, we can assume, a valuable way to live – a valuable type of relationship to have.¹² But in order to relate to someone as an equal one must refrain from certain kinds of inquiry, specifically inquiries about the variations amongst people with regard to these fundamentally important properties. Such inquiries will undermine the basis of equality as long as the features they uncover are allowed to figure in determining how the person ought to be treated (as it seems they must be, since they go towards qualifying the person for basic respect). In which case we must avert our gaze from such features, abstaining from certain forms of knowledge.

Now this kind of view faces an immediate objection, which may have been troubling the reader from the outset. It is impossible for us to do away with knowledge that we already have. We are condemned to know what we know, and cannot decide not to know it. The only question, it might be said, is how we can use the knowledge wisely.

Carter, however, enters an important qualification in response to this criticism. It is not quite that respect requires one not to *know* certain things; how could it, after all, since what one knows is not under one's voluntary control? Rather it

¹⁰ Carter, 'Respect,' p. 550.

¹¹ Carter, 'Respect,' p. 551.

¹² E. Anderson, 'What Is the Point of Equality,' *Ethics* 109 (1999), pp. 287-337. I take it that Carter himself is attempting to establish this claim; we need only assume it here.

requires one not to *act on* that knowledge: ‘when we do perceive people’s varying agential capacities we refuse to let such perceptions count as among the reasons motivating our treatment of those people.’¹³ Respect, therefore, has something in common with Raz’s well-known treatment of exclusionary reasons in authority and elsewhere.¹⁴ Exclusionary reasons are second-order reasons: reasons to do with other reasons. Where there are exclusionary reasons, they consist in reasons not to take other reasons that would normally be relevant into account in our deliberations. To take an example from the realm of authority, my daughter may find that various competing considerations bear on the question how she should spend the next ten minutes; but when I tell her to do the dishes, if I have authority over her, the directive replaces those other reasons. It is not, on Raz’s account, that she has *stronger* reason either to do the dishes, or to do as I say, than she has to do anything else, and hence that the directive *outweighs* those other reasons; it is rather that the authoritative issuing of the directive *precludes* her from attending to the relative strength of the various reasons. It asks her to exclude those considerations from her practical deliberations. Similarly, respect asks that, even though we may know something about how two people compare with respect to rationality, as long as we also know that they are both rational enough, we put those variations out of practical consideration, treating them externally, formally, as equals.

The virtue of respect in this way structurally resembles the virtue of tact. Tact does not involve not knowing; rather it is a virtue activated when one is *in possession of* knowledge, or at least in possession of an opportunity to gain some knowledge. Tact is the virtue of knowing when and when not to inquire, when and when not to treat such knowledge as a basis for action, or for changing the way one treats someone. Tact is a virtue because good relations sometimes require such reticence. This manoeuvre is hardly alien to human moral psychology. It might be hard to imagine a successful form of human social life

¹³ Carter, ‘Respect,’ p. 551.

¹⁴ J. Raz, *The Authority of Law* (Oxford: Oxford University Press, 1979), Ch. 1. On exclusionary reasons in promising, see ‘Promises and Obligations’ in P. M. S. Hacker and J. Raz (eds), *Law, Morality and Society* (Oxford: Oxford University Press, 1977); D. Owens, *Shaping the Normative Landscape* (Oxford: Oxford University Press, 2012).

that does not value a capacity for tact or discretion, and hence epistemic abstinence in some form or another.¹⁵

4. Relations of respect

We can now see that Carter's view contains an account, not just of what the *basis* of basic respect is, but of the *kind of treatment* that respect requires. Carter's account thus differs from the way in which similar considerations might be deployed in an argument about the basis of moral considerability. We can see the difference by setting out a version of the latter. Say that moral considerability involves qualification for the possession of equal basic moral rights. Someone might argue that, in determining which beings are morally considerable, we should determine the relevant threshold, and hold that differences above that threshold will be morally irrelevant as far as the possession of those basic rights goes. With respect to differences above the threshold, we should therefore adopt an attitude of evaluative abstinence; such differences will not affect a person's possession of equal basic rights. This argument gives epistemic abstinence a role. But it limits the role of epistemic abstinence to that of determining who does and who does not possess basic rights. It does not claim that one of forms of treatment to which we have a right is not to be treated in a way that depends on making inquiries into one's level of possession of those qualities qualifying one for basic respect. However, Carter's position does make this further claim. It tells us not just who to respect, but something about how to do it; specifically, it tells us that respecting a person involves refraining from forms of treatment that are based on such inquiries. This is what makes it possible to derive from Carter's account the claim that respect requires maintaining a certain kind of distance.

Nevertheless, I now want to consider another potential objection to Carter's account: that it cannot account for relations of respect. The concern here is that Carter's account may provide us with an answer to the question of who counts for the purposes of equal respect, and may show how, by deploying opacity, the

¹⁵ On this point, see G. Simmel, 'The Sociology of Secrecy and of Secret Societies,' *American Journal of Sociology* vol. 11 (1906), pp. 441-498; T. Nagel, 'Concealment and Exposure,' *Philosophy and Public Affairs* vol. 27 (1998), pp. 3-30.

distribution of basic goods and freedoms equally need not be disrupted by the fact that people possess respect-worthy capacities to differing degrees; but that Carter has not yet articulated how to move from this essentially static conception of respect to the more dynamic notion of relating to another as an equal. The reason for this, the objection might go, is that relating to another as an equal in an on-going way involves a certain kind of responsiveness to that other – responsiveness to at least some things about them. But which things about them are and are not permitted as a basis for opacity respect? Carter’s view as presently articulated does not explicitly address this question: it might be said, for the reason that it is largely a negative conception of respect. Positively it provides us with a threshold that acts as a qualifying criterion for respect-worthiness. And it thereby provides us a basis for distributive equality. But thereafter it only tells us what we must *not* take into account if we are to treat another person as an equal. It does not offer a positive account of the responsiveness to another person that is constitutive of relating to them with respect, which would tell us which of their features we are and are not permitted to respond to in treating them as an equal. An account of such responsiveness would have to say something about what we are responding to, and how we are appropriately to respond: what respectful responses consist in. Carter’s notion of opacity-respect cannot capture these intuitions,

Now it may be correct to say that Carter has not spelled out the implications of his view for the ‘dynamic’ demands of relating to another as an equal in an on-going way. Nevertheless, it seems that his view could be extended naturally enough in a way that captures the intuition that relating to another as an equal involves some disposition to take them seriously, to give them a hearing, to invest them with *prima facie* credibility, and so on.¹⁶ Let us assume that the capacities that qualify a person for respect have to do with rational and moral deliberation and agency. A person, that is, qualifies for equal basic respect by virtue of meeting a threshold ability to think intelligently for themselves and govern their action accordingly. If this is the basis for respect then, Carter might

¹⁶ That these features are aspects of equal relations is an implication of the account of epistemic justice presented in M. Fricker, *Epistemic Injustice: Power and the Politics of Knowing* (Oxford: Oxford University Press, 2007).

say, relating to another as an equal has to involve treating them always in such a way that their possession of those agential capacities is most salient. What will that involve? I think the decisive thing is that it will involve treating what a person thinks and does as coming about in the presence of that threshold level of agential capacities. In other words, it involves interpreting what they do – for practical purposes – as a deliverance of (at least) that threshold level of thought and agency. Of course, human behaviour has many causes. But to treat people as equals, the thought here is, is to treat them equally on the assumption that their behaviour involves input from at least the basic threshold level of agential capacity.

This way of interpreting human agents as having an equal set of basic capacities grounds some practical assumptions about their possession of credibility, authority and competence that underpin the demand that each should be ‘taken seriously’ in practical life. It means that, up to a certain point, all are to be treated as having the same abilities to think for themselves and to govern their behaviour. They can therefore be included as equals in characteristic forms of human interaction such as conversation, inquiry, forms of play, friendship, collegiality, and so on. But it also means that people can appropriately be subject to certain kinds of criticism, for instance, if they do not deploy their capacities in the way that someone with the threshold level can reasonably be expected to do. In short, on the basis of making practical assumptions about their possession of agential capacity, we can ground the demand that we include others equally in forms of interaction that are structured by expectations that we have of one another and hold one another to.

However, in order to maintain this way of relating to one another as equals we require opacity: the fact that we share a threshold level of capacity has to be more salient, for the purposes of these relations, than are our differences. We are uninterested, for the purposes of relating as equals, in differences in capacity over that threshold: we are simply interested in interpreting behaviour as a deliverance of a threshold level of capacity for self-government. We include people in forms of human interaction, hold them to expectations and subject

their views and actions to criticism on the assumption that they are equals, and without asking them to prove their credentials. Of course, we can always 'look within' to see whether one person had greater aptitude for self-government than another, and whether one faced greater obstacles to meeting those expectations than another. But it is not straightforwardly true that the knowledge we might thereby gain can be treated as relevant if we are to maintain our relations as equals. When relating to one another as an equal, our default position is to take at face value the deliverances of the person's thought and agency, and not to look at the 'mess inside' from which it stemmed.¹⁷ We treat the person as an equal by dealing with them in the image of a rational agent presenting herself to the world on the basis of rational deliberation. We take at face value this rational self-presentation, averting our gaze from its psychological sources.

5. Respect and neuro-intervention

Now even if I am correct about the implications of respect, it remains true, as I said at the outset, that we cannot draw any simple conclusions about how offenders should be treated. Considerations of respect are weighty, and, if the argument here is correct, exclude giving practical relevance to knowledge that would otherwise call to be taken into account. However, respect is not the only value, and I doubt that it has absolute priority. This means that there is scope for dilemmatic situations in which sentencers have to recognise non-excluded considerations favouring incompatible courses of action. With that qualification in mind, though, we can say that my interpretation of Carter's view suggests what we might call a *minimal* and *external* picture of criminal justice.

What Carter's proposal does is to suggest a link between respect and keeping one's distance. It claims that respect requires abstaining from certain kinds of assessment of persons and taking them rather as they present themselves.

I will follow those who think that the criminal sanction has an expressive or communicative element, and who therefore conclude that liability to that

¹⁷ This is not in any way to deny that there is such a mess, and that its study is an important part of self-understanding. The point here is rather about the circumstances in which such study can and cannot be treated as practically relevant. See P. Goldie, *The Mess Inside: Narrative, Emotion and the Mind* (Oxford: Oxford University Press, 2012).

sanction, analogously to liability to moral criticism, depends on the possession of certain capacities of basic moral agency – where the possession of those capacities makes it fair to hold a person to certain normative expectations and subject them to a range of responses for violations.¹⁸ Thus moral criticism is compatible with respect in a way that other responses that depend on giving practical relevance to factors other than the offender’s rational self-presentation are not. If moral criticism is the basic purpose of the criminal sanctions then, once we have assured ourselves that an agent meets the basic conditions of liability, respect requires that we treat her, from the practical point of view, as opaque. For anything else – any attempt to ask why the person acted as they did or how they could be prevented from acting in that way in the future – would involve inquiring into differences in moral capacity; whereas respect insists that it is not the differences in possession of some scalar property, but rather the equal possession of a range property, that should be salient in our practical dealings with fellow equals. And offenders, like all those who meet the threshold condition, have to be treated with basic respect. (Though they may lose or forfeit other rights, they do not lose the right to basic respect.)

Opacity respect, then, is compatible with holding a person to account for her actions, but not compatible with subjecting her or her motivations to the scrutiny that would bring to light considerations that might otherwise be reasons for taking *other* kinds of action against her – for instance, giving practical weight to any knowledge we have of the offender’s underlying psychology, even if this knowledge might let us see how that offender’s behaviour might best be altered in the future. The argument from opacity respect suggests, then, that sentencing should revolve around non-individualised responses based on the category of the offence rather than the aim of behaviour-modification. The conclusion of the argument is therefore that there should be a strong presumption against neuro-interventions. This conclusion does not rely,

¹⁸ For instance, if the justification of the criminal sanction gives a central place to moral communication (to the offender) then basic moral agency (for instance, an ability to understand the wrongness of one’s action and take responsibility for it) will be a necessary condition for liability. It is not clear that basic moral agency should be necessary for liability if the point of the criminal sanction is simply deterrence.

however, on the assumption that coercion or setbacks to our interests in physical or mental integrity are sufficiently bad to ground such a presumption. It rests rather on a view of what is required if we are to relate to one another, person to person, as equals. Chemical castration is a violation of basic respect – not because of its coercive or intrusive quality as such, but because it involves making inquiries that undermine that potential offender’s status as an equal.

In previous work, I have attempted, in Strawsonian fashion, to link such epistemic abstinence, not only to relations of equality, but to the perspective we adopt when we deal with someone as a participant in ‘ordinary adult interpersonal relations’ more generally.¹⁹ Strawson contrasts this ‘engaged’ or ‘participant’ perspective with what he calls ‘the objective perspective.’ In the latter, we deal with someone rather as a causal system in which various inputs cause a potentially predictable range of outputs, and where the inputs can in principle be manipulated to lead to beneficial outputs. Strawson’s argument is intended as a vindication of the interpersonal perspective, but it proceeds without an explicit account of the normative commitments of that perspective, or an explicit argument as to what would be lost if we only ever employed the objective perspective. In order to make Strawson’s conclusions more compelling, I have argued that ‘ordinary adult interpersonal relations’ should really be seen as an inherently valuable form of interaction. Part of their value lies in the value of being included in collaborative interactions in which one is treated as an agent capable of a kind of independent self-governance. Independent self-governance has to do with possession of a capacity for competent performance in some social practice or interaction without the need for constant surveillance or supervision from others. My understanding of such competence is that it is a range property, in the sense that possession of competence above a given threshold qualifies one for non-scalar respect: it earns one a kind of *authority* and *standing* within the interaction.²⁰ Such standing consists in a right to be

¹⁹ P. F. Strawson, ‘Freedom and Resentment,’ *Proceedings of the British Academy* 48 (1962), pp. 1-25. For my interpretation of these arguments, see *The Apology Ritual: A Philosophical Theory of Punishment* (Cambridge: Cambridge University Press, 2008), Ch. 3.

²⁰ For the relevant notion of authority, see P. Pettit and M. Smith, ‘Freedom in Belief and Desire,’ *Journal of Philosophy* 93 (1996), pp. 429-449. For an example of disrespectful behavior consisting

taken seriously, to be given pro tanto credibility, and to be included in the practice on terms that reflect one's achievement of a threshold level of competence. It consists, furthermore, in a right to the space necessary to determine one's own contribution to the cooperative interaction that the practice involves. Having authority as a collaborator means that it is up to you to decide the nature of your contribution; you can then be subject to reasonable expectations regarding competence and goodwill, and to reactive attitudes should those expectations be violated; but it is normally inappropriate and contrary to your authority for another to pre-empt your contribution, or to manipulate you in such a way that your contribution is no longer an expression of your own considered and free competent response to the situation. Giving someone space to determine their own contribution requires evaluative abstinence: it requires refraining – up to a point at least – from inquiries about whether their contribution is likely to be a good one. Respecting their authority involves allowing them the right to determine their contribution, even though this can sometimes lead to inefficiencies and worse. If what makes certain practices non-instrumentally valuable is in part their character as free collaborations between self-governing agents, the value of the practice rests on participants maintaining a kind of distance that allows each to determine the character of his or her contribution. It requires that certain aspects of our collaborators are, from a practical point of view, opaque.

Indeed, the insight grounding opacity might be expressed in other terms. For instance, it might be put in terms of trust. Trust is sometimes said to be based on an assessment of a person's trustworthiness, or of the likelihood that that person will take one's relying on them as a reason to act in one's interests.²¹ However, looking at trust from the perspective of opacity, one can see the possibility of another position. For it might be said that by the time one is making a reasoned assessment of a person's trustworthiness one has *already* lost trust in them. Trust, the thought might be, is implicit. Trust might then be said to exist when it doesn't occur to one to ask whether another person is trustworthy, or where one

in the failure to acknowledge such authority, see S. Shiffrin, 'Paternalism, Unconscionability Doctrine and Accommodation,' *Philosophy and Public Affairs* 29 (2000), pp. 205-250.

²¹ R. Hardin, *Trust and Trustworthiness* (New York: Russell Sage Foundation, 2002).

refuses to inquire into the matter. Reasons for trust would therefore be exclusionary second-order reasons; they would be reasons to rely on others but not to do so on the basis of a calculation of the wisdom of doing so. Trust would involve opacity since it would involve a refusal to inquire. Furthermore, if Carter's arguments are correct, some form of trust would be required by basic respect. Up to a point (and I have been careful not to attempt to specify this point), we owe potential offenders (even those who have offended in the past) respect and trust, and that will normally rule out the kinds of intrusive inquiries that would be needed to inform neuro-intervention.

6. Objections and replies

In this section I would like consider some objections to the argument we have sketched out so far. In particular I am concerned to rebut the objection that Carter's proposal proves too much. For instance, Carter's claim is presumably that opacity respect is a perspective that we are required to adopt in certain circumstances or contexts, but not all. Otherwise it would rule out intimate relationships, since one of the things distinctive about intimate relationships is that we are precisely allowed to, and expected to, take a concern with another person's inner life, how things are going with them, and what support they might need. Furthermore it would rule out educative relationships – at least where these hold between adults – since again it is essential to such relationships that the educator inquire as to the student's current level of understanding, their capacities and potential, in order to ascertain the best pedagogical approach. Finally, it would rule out relations of psychological help, from counselling to CBT to psychoanalysis, which again involve the subject in a certain amount of self-revelation. Assuming that these relationships are desirable, or at the very least permissible, Carter's claim might seem too strong.

I think three lines of response are possible here; although I do not commit myself to any particular one, let me canvass them briefly. The first says that the requirement of basic respect is always in force, in any situation, but that it is open to agents to waive their right to opacity and to 'let others in.' The basic idea behind respect, on Carter's view, is something like Dworkin's 'swimmers in

separate lanes' analogy.²² This sees us as morally separate, responsible primarily for ourselves and with only limited responsibilities for others; and having rights to exclude others and determine our own path (consistent with the same rights for others). Nevertheless we are not condemned to solitude; we can make connections with others and 'allow them in' in various ways. This, it might be said, is what we do in intimate relationships, relations with psychological and health professionals, educators and so on: as long as we give our full and free consent to such arrangements, they are quite compatible with the separateness of persons; the scrutiny that they involve, which would otherwise be intrusive and disrespectful, might indeed be deeply rewarding. Hence it is not the nature of the scrutiny by itself that determines its moral character; rather it is the mode of relation in which the scrutiny takes place that determines whether it is disrespectful or not.

The second line of response takes a more context-relative approach. It takes seriously the thought that there are contexts in which the requirement of basic respect – at least insofar as it involves the demand that we relate to one another as equals – is not in force, and in which it would be inappropriate to insist on it, but where this is not best explained by the mechanism of consent and the waiving of one's right to respect. For instance, it is sometimes said that there is no justice in the family – which we might interpret to mean that, even when family members are adults, the family is a context in which the force of norms of equal status, respect, distance and separateness is lessened. Family members might have a duty to support one another precisely with the maintenance of the kinds of agential capacities that qualify us for basic respect (but which it might be hard to maintain in the absence of such support: in this sense, the possibility of independence rests on the reality of dependence) even when that support is not asked for or consented to. The same might go for some other contexts: in the context of certain sorts of relationships of friendships, it might be inappropriate to insist on the opacity that, according to Carter, marks basic respect. If it were possible to come to be in a relationship of friendship, or another sort of support

²² R. Dworkin, *Justice For Hedgehogs* (Cambridge, Mass: Belknap Press, 2011). Note that Dworkin's analogy is meant to explicate the limited nature of our responsibility for others.

relationship, without having given full consent (though what this would involve will depend on what one takes consent to consist in – for instance, whether it can be tacit), then another person might be within their rights in not treating you as opaque. Nevertheless, it might be that Carter’s thought about basic respect should regulate our dealings with one another outside of the context of certain special relationships.

Thirdly, we might claim that, whatever is true of other forms of relationship, respect should characterise the relations between the state and its citizens, and hence that Dworkin’s analogy represents an ideal that should regulate that domain, regardless of whether the basic respect is in force in every moral context. The first two responses can be seen as responses to the concern that Carter’s position makes a fetish of independence, and that it would make it morally impossible to acknowledge the reality of the ways in which even adults are mutually interdependent. These first two responses represent ways to show how a social form that values independence might be compatible with support for dependence. This third response, by contrast, says that there is a particular domain in which agents should be treated under the idea of independent self-governance, and that is when they are treated as citizens. It is an important part of good political relations, the thought is, that the political community be seen as an association of independent co-operators, rather than as some sort of morally undifferentiated social aggregate: this is a Rawlsian thought about the ‘separateness of persons.’ The importance of respect, of treating citizens as free and equal, seems to be a fairly fixed point in liberal political philosophy: this argument would claim that coercive neuro-intervention is incompatible with equal respect for citizens, and that when we unpack the notion of respect it turns out to involve opacity. If successful, this gives us an argument against neuro-intervention imposed by the state that is grounded, not in concerns that such interventions are ‘unnatural,’ or libertarian views about self-ownership, or worries about the pernicious effects of pharmaceutical interventions on the authenticity of the individual, but rather in claims about the limits to the state’s intervention into individual lives if the state is to meet standards of respect.

However, Carter's view might prove too much in another way. Would it rule out commonly accepted features of criminal justice in particular? This concern might arise in two different ways. First of all, it might be said that criminal investigations involve intrusive inquiries into what an agent has done, and in particular into whether they have acted intentionally, accidentally, or under psychological compulsion, and that treating agents as opaque rules out the application of defences like duress, loss of control, automatism and so on. Secondly, it might be pointed out that both at sentencing and at parole hearings, official attention is paid to issues of public safety and to individual rehabilitation, both of which can involve some attempt to investigate and predict how a person will act if they are left to their own devices, and what action can be taken now to alter things for them so that they do not act in that way. Neuro-interventions are one – perhaps extreme – version of public safety/rehabilitation interventions. But will the argument for opacity respect rule out even interventions that many might regard as more innocuous?

Let me deal with these objections in turn. A quick answer to the question about defences would be that defences are put forward voluntarily by defendants at trials, and can therefore be seen as an expression of consent to have their behaviour scrutinised in a way that would not normally be permitted. Since we have already argued that interventions that are consented to are not incompatible with opacity respect, it would follow that allowing someone to enter a defence that then authorises an investigation into the extent of their agential capacities would not be incompatible with opacity respect.

However, some might think this response too quick. After all, sometimes pre-trial investigations may need to be intrusive in order to determine whether a given suspect has a case to answer. In response, then, three points need to be made. First, being held accountable for one's actions is not incompatible with basic respect; if anything, it is an expression of basic respect that one's actions are publicly acknowledged as one's own, performed for reasons for which one might have to answer. Secondly, while there are various senses of 'doing,' the sense relevant to liability is the sense in which what one does is that for which

one is answerable, and it is not in general possible to say what a person is answerable for without some attempt to determine what their intentions were, and what their reasons were for acting as they did. Therefore capturing the character of the wrong involved, say, in a particular case of intentional killing of another person will require an investigation aimed at capturing the light in which the agent saw what they were doing, and hence the normative reasons on which they acted, in order to assess whether they have any good justification for acting as they did. Trying to reconstruct a person's reasons for acting as they did involves an attempt to put oneself in their place, and hence involves some sort of inquiry that 'gets into their head.' Yet rational explanation of action is surely not intrusive as such: indeed, as we saw earlier, it affirms the basic possession of agential capacities by interpreting the agent's behaviour in their light. Thirdly, however, the argument of this paper should not be interpreted as the claim that intrusive intervention as such – or being subjected to scrutiny as such – is disrespectful. That would indeed be too simple and blunt a way to interpret Carter's account. Rather such intervention is problematic when it involves a failure of basic respect; specifically, when it involves an inquiry into the extent to which one possesses, and is likely to exercise, capacities above the threshold necessary to qualify for basic respect. Inquiries that are aimed at discovering how one acted and what one is answerable for are not inquiries that are intrusive in the sense that violates the demands of opacity-respect. For instance, an inquiry into one's motivations to see, say, whether one had a justification defence is simply an investigation into the question of what one has done, what can be attributed to one's agency in the sense that one can be expected to answer for it. But even the application of excusing defences – which I take to concern, not whether what one did was wrong, but whether in a full sense one can be said to have done the thing in question – is quite compatible with opacity respect. Granted, in order to say whether one acted accidentally, or in ignorance, or under the influence of medication that removed one's rational control, we need to look at one's agential capacities. But we do so in the service of determining the nature of one's exercise of one's own agency; it involves taking the exercise of one's agency at face value, not second-guessing or pre-empting it. Investigating the nature of one's action in order to ensure that one is treated in ways that

reflect one's exercise of one's own agency respects rather than disregarding the qualifying conditions for opacity respect. Therefore although criminal investigations are intrusive in an important sense, they are not (necessarily) intrusive in the way that violates opacity respect, since they are compatible with the idea that a person is defined by how they will to present themselves, and that 'the mess inside' that issues in such action should be treated as opaque.

Therefore I conclude that arguing for opacity respect does not have the counter-intuitive implication that criminal investigations, or criminal defences, are disrespectful. What of the second aspect of this criticism, that opacity respect would rule out many seemingly innocuous public safety and rehabilitative interventions? The most global, and hence threatening, version of this criticism points out that there are many instances in which private individuals and state officials deploy psychological insights in order to predict behaviour and take action accordingly. The less ambitious version of this criticism claims that many post-conviction interventions, including sentences such as Drug Rehabilitation Orders, or required anger-management courses, as well as parole hearings and other assessments that seek to quantify the risk that the offender poses to the public. Both of these criticisms, however, point to the fact that we often, and apparently innocuously, deal with people on the basis of the probability they represent of causing some desirable or undesirable state of affairs, rather than as individuals who must be treated as opaque beyond their possession of basic agency. Would this be incompatible with Carter's recommendation of epistemic abstinence?

These criticisms require more extended treatment than I can give them here, but let me make three points quickly. Firstly, not all of these ways of seeing or treating people involve imposing something on them in the way that neuro-interventions would be imposed. Using someone's behaviour as data for predictive purposes might be perfectly compatible with opacity respect as long as it doesn't lead to their being dealt with disrespectfully by having treatment imposed on them against their will. Behaviour modification treatment may be quite acceptable if it is given consensually. And predicting someone's behaviour

need not be unacceptable in itself, on Carter's terms.²³ Secondly, not all of the 'behaviour orders' that are used as sentences subvert or pre-empt the offender's rational capacities in the way that chemical castration would. Rather they often engage reason and challenge the offender to think differently. It is not clear that interventions that operate at this rational, personal level, rather than the neurological level, raise the kinds of concerns that I have been expressing here. Thirdly, however, I think we should indeed countenance the view that, where an offender has a sentence imposed on her that requires her to submit to some neurological intervention against her will, that this violates her status as a self-governing agent who has the agential capacities necessary for basic respect. It pre-empts her responsibility to do this thinking for herself. I do recognise that offending behaviour, especially when persistent, changes the offender's normative situation, and makes her liable to treatment that she was not liable to beforehand. However, the kinds of treatment to which she is liable – even punishment – can and should be meted out in ways that recognise the status to which her basic agency qualifies her, and which therefore do not renounce the ways of interacting with human beings that those capacities make possible.

7. Conclusion

In this paper I have explored a strategy for arguing against the coerced enhancement of serious offenders. This strategy does not appeal to the unnaturalness of enhancement but rather to its intrusiveness. It argues that what is wrong with such enhancement is the scrutiny on which it is based. Such scrutiny, according to this strategy, is incompatible with the opacity required by respect. This helps to explain why the aspiration to inclusiveness in human relations can require a commitment to epistemic abstinence. Unbridled deployment of the knowledge modern science delivers can undermine valuable forms of human interaction.²⁴

²³ For some discussion of these issues from a perspective different from that adopted here, see F. Picinali, 'Base-Rates of Negative Traits: Instructions for Use in Criminal Trials,' *Journal of Applied Philosophy* 33 (2016), pp. 69-87.

²⁴ For discussion of an earlier version of this paper, I am very grateful to participants at the 'Treatment for Crime' workshop held at the University of Oxford in April 2015. I am particularly grateful for comments and questions from Ian Carter, Matt Matravers, Tom Douglas, David Birks, Kasper Lippert-Rasmussen, Emily McTernan and Matthew Clayton.