

Vincent, Nicole A (2010) *Madness, Badness and Neuroimaging-Based Responsibility Assessments*, in Michael Freeman (ed) *Law and Neuroscience, Current Legal Issues (vol 13)*, OUP.

MADNESS, BADNESS AND NEUROIMAGING-BASED RESPONSIBILITY ASSESSMENTS

NICOLE A VINCENT – N.A.VINCENT@TUDELFT.NL

On the capacitarian account responsibility tracks capacity — i.e. reductions in a person’s agency-relevant mental capacities entail reductions in the degree of their responsibility. This entails that if the mental capacity to (e.g.) empathise was necessary for moral agency, then a person that commits murder but who lacks this capacity should be at least partially excused. The capacitarian account also allows that if neuroimaging scans could reveal people’s mental capacities then such scans could play some role in assessing responsibility. However, viewed from another angle, a person who lacks the capacity to empathise could also be described as cold blooded killer – i.e. as a bad rather than a mad person – and under this description it seems more appropriate to condemn rather than to excuse them. Although this might seem to challenge the capacitarian account, and with it the idea that neuroscience might have some role to play in helping us to assess people’s responsibility, I will argue that this challenge is illusory. On my account, this challenge rests on a conflation of two different kinds of culpability, and on a failure to notice that each of these kinds of culpability plays a role at a different stage in the criminal trial.

Suppose that neuroimaging scans could reveal people’s mental capacities, and that a scan of Conrad’s brain – a man with a history of violence who has been convicted of murder and diagnosed as a psychopath – revealed that he lacks the capacity to empathise. Precisely what, if anything, would such evidence tell us about Conrad’s responsibility; would it *excuse* him (even if only partially), or would it *condemn* him?

In my view, such evidence would at least partially excuse Conrad since it would show that he lacks a mental capacity which is arguably required for moral agency — i.e. on my account such evidence would testify for the *defence*.^{1 2} However, on Heidi

¹ Nicole A Vincent, ‘Responsibility, Dysfunction and Capacity’ (2008) 1 *Neuroethics* 199-204.

² Nicole A Vincent, ‘Neuroimaging and Responsibility Assessments’ (2009) *Neuroethics*.

Maibom's account, such evidence would condemn Conrad even further since she thinks that it would graphically show (*via* images of his brain) that he is a bad person all the way down to the core — i.e. on her account such evidence would testify for the *prosecution*.³ So who is right, Maibom or me — would such evidence excuse or would it condemn? Marga Reimer thinks that neither of these two positions is exclusively right, claiming instead that such evidence *both* excuses *and* condemns. Reimer writes that “both defence and prosecuting attorneys might make use of the relevant brain images. Indeed, they might appeal to *the very same* images, but characterize them in radically different ways — brains that are ‘damaged’ vs. brains that are ‘evil’”.⁴

If Reimer was right then such evidence would be a lot less useful to the defence than I have suggested. However, I will argue that once we clearly distinguish condemnation of people *for who they are* from condemnation of people *for what they do*, and once we realize that each of these two kinds of condemnation plays a role at a *different* stage in a criminal trial, we will see that at the guilt determination stage such evidence clearly favours the defence. In what follows, after offering an account of how a future neuroscience might one day play a role in helping us to assess people's responsibility, I will then explain Reimer's and Maibom's concerns and relate them to this account, and I will end by rejecting Reimer's claim that such evidence *both* exculpates and condemns and Maibom's claim that such evidence condemns *rather than* excuses.

1. NEUROSCIENCE AND RESPONSIBILITY ASSESSMENTS

The point of this section is to explain how a future neuroscience might play a role in helping the criminal law to assess people's responsibility. I will begin by describing a “murder mystery” example; I will then explain how an assessment of responsibility would normally be conducted in such a case; and finally, I will explain what neuroscience might one day add to the process of assessing responsibility in cases like

³ Heidi L. Maibom, ‘The Mad, the Bad, and the Psychopath’ (2008) 1 *Neuroethics* 167-84.

⁴ Marga Reimer, ‘Psychopathy Without (the Language of) Disorder’ (2008) 1 *Neuroethics* 185-98, p. 196, emphasis added.

this. Since these ideas are discussed at length elsewhere,⁵ and the main point of the present paper is not to develop these ideas but to defend them from a novel objection, I will therefore keep this discussion brief, although references will be provided to sections of that other paper where greater detail and citations to relevant literature can be found.

1.1. MURDER MYSTERY

Imagine that you stumble upon Smith's dead body while strolling through a forest. It seems that Smith died of a gunshot wound to the head, and that his body was then hastily concealed beneath the bush where you found him. What processes might we engage in to determine who is responsible for Smith's death?

1.2. ASSESSING RESPONSIBILITY *WITHOUT* THE HELP OF NEUROSCIENCE⁶

Our first question will probably be "Who dunnit?", and so in the beginning our inquiry will involve such things as: gathering up witnesses and suspects; perhaps conducting a line-up where eye witnesses can identify prime suspects; the police might then interrogate various people; in court, both witnesses and suspects might eventually be questioned; and physical evidence such as finger prints, spent bullet cartridges, DNA samples and so on might also be collected, examined and presented in court. Many epistemic barriers stand in the way of answering the *who dunnit* question, but once these pieces of the puzzle are put together we may discover that Jones is the one who shot Smith dead. Very roughly, in criminal trials the *actus reus* inquiry addresses the *who dunnit* question.

But to establish that Jones is responsible for Smith's death, we need to show a lot more than just that she "dunnit". Rather, given causal indeterminacy – i.e. the fact that any outcome is a result of many causal contributions – we must also show that her causal contributions were of particular significance, and the way that the criminal law does this is by conducting a *mens rea* inquiry to determine whether the party who committed the prohibited act was at fault, guilty or culpable. To see how this might work, consider a sample of defences that might be offered by Jones to relieve herself

⁵ Nicole A Vincent, 'On the Relevance of Neuroscience to Criminal Responsibility' (forthcoming) *Criminal Law and Philosophy*.

⁶ I discuss this in detail in N. 5, above, §2.

of responsibility for Smith's death despite her obviously-pertinent causal contribution. For instance, Jones might point out that she shot Smith in *self-defence* when he ambushed her on her stroll through the forest, and as long as what she did is viewed as a reasonable response – e.g. not an unwarranted use of extreme force – then this should suffice to establish that Smith was at fault for his own death (because he should not have attacked her), and thus Jones' causal contribution would not translate into a finding that she was responsible for his death. Alternatively, suppose that Jones was out hunting in a well-known, sign-posted and cordoned-off area of the woods which was only supposed to have other hunters in it, all of whom wore brightly coloured clothing and who knew each others' locations, and that there was simply no way of her knowing (nor any reason to suspect) that Smith – a thrill-seeking prankster who liked to frustrate hunters by hiding in bushes and scaring away their game – was hiding in those bushes. In such a case it is again likely that Smith's rather than Jones' actions will be viewed as the salient causes of his own demise, and hence that he would again be picked out as the person who is responsible for his own death, because Jones acted *reasonably* (her actions were not unduly risky) whereas Smith did not (his actions were too risky).⁷ Furthermore, to the extent that *intentions* affect culpability, given that Jones did not even realize nor had any reason to suspect that Smith was in that bush, she certainly did not shoot him intentionally (and ex-hypothesi she was not negligent), and so she can't be culpable on account of having culpable intentions either. Finally, Jones' culpability may also be reduced by citing psychological considerations — for instance, that she is insane; that she was drugged, dazed and confused; or that she is a child. These and other *mens rea* defences offer reasons (excuses or justifications) to suppose that in the final analysis the party in question should not be blamed for how they acted, and hence that they are not responsible for the outcome despite their causal involvement.

However, what is specific to how the cited *psychological* defences work in the context of the *mens rea* inquiry,⁸ is that culpability is reduced by showing that the

⁷ My point is not that we would necessarily make this substantive evaluation about the risks involved, but rather that *if* we made that evaluation *then* Smith would be deemed responsible.

⁸ Not all psychological defences address the *mens rea* inquiry. For instance, the “automatism” defence – e.g. that the actions were performed while sleep-walking – defeats an accusation of responsibility, by appeal to the idea that true actions (unlike mere body movements) must be

accused person lacked a mental capacity which is necessary for acting in the prescribed way, and schematically the law's reasoning here looks something like this:

mental capacity → can & can't → duties → culpability

From right to left this schematic reads: one is culpable when one violates one's duties;⁹ one's duties depend at least in part on what one can and can't do;¹⁰ what one can and can't do depends at least in part on one's mental capacities;¹¹ and so the scope and degree of one's culpability varies among other things in proportion with

voluntary — i.e. the automatism defence addresses the *actus reus* inquiry. For discussion of violence related to sleep disorders see Mark R. Pressman, 'Disorders of Arousal From Sleep and Violent Behavior: The Role of Physical Contact and Proximity' (2007) 30 *SLEEP* 1039-47.

⁹ I take this to be an analytic truth, though I use the term "duty" in a broad sense that is meant to capture the many different things all of which we ought to do.

¹⁰ There are two ways of conceiving the relationship between capacity, duty and blame. On one account, capacity plays a *positive* role by *generating* duties: we ought to do what we have most reason to do, and what we can and can't do (presumably along with a range of many other things) generates the reasons that we have to do various things. On this first account, if I can not save a child from drowning — perhaps because I do not know that they are drowning, or because I can not swim, or because I do not have a rope to throw to them — then in the final analysis it is simply not true that I ought to save them (unless I am responsible for the fact that I can not do this); thus the reason why I would not be blameworthy for not saving them is because I did not, in the final analysis, actually have that saving duty. On another account though, capacity plays a *negative* role by *regulating* duties: regardless of the source of our duties, on this second view our *incapacities* can justify departures from those duties. On this latter account, the three cited considerations do not outweigh the saving duty, but rather they justify departing from it — i.e. the reason why I would not be blameworthy is because although I did have the saving duty, my incapacity provided an excuse for departing from it. Two advantages of the second view is that only it has an explicit place for excuses (which occupy an important place in legal discourse), and it arguably also more adequately captures the rich structure of practical reasoning in which some considerations *discount*, *undermine* and *invalidate* (rather than just *outweighing*) other considerations. Never the less, I suspect that both views of the relationship between capacity, duty and blame generate the same conclusions about when a person is to blame and when they are not to blame, and since I find it simpler to explain ideas within the conceptual framework of the first account, in what follows I will therefore reason within that framework.

¹¹ Though we arguably have greater control over our *internal* constraints than over *external* constraints, this does not negate the fact that internal constraints can prevent us from being able to do various things.

one's mental capacities. Put a different way, one is culpable when one acts contrary to how one ought to act, and one important determinant of how we ought to act – or perhaps a consideration that *undermines* claims about how we ought to act¹² – is what we *can* and *can not* do. But since some of the things that we can and can not do relate to our mental capacities, a person who non-culpably¹³ lacks a mental capacity that is needed to act in some prescribed manner may therefore cite this incapacity or reduced capacity as an excuse for failing to act as prescribed. The cited psychological defences therefore reveal the law's *capacitarianism* — i.e. the underlying idea that among other things *responsibility tracks capacity*.

In law, as in morality, responsibility is partly a matter of whether the party in question is culpable – i.e. whether they acted contrary to how they should have acted – and partly a matter of their causal involvement, and this is why we must conduct a *mens rea* inquiry as well as an *actus reus* inquiry.

1.3. HOW A FUTURE NEUROSCIENCE MIGHT FIT IN¹⁴

A future neuroscience – i.e. one that overcomes the limitations of current science and technology – might fit into the above picture in the following ways.

Firstly, neuroscience might help us to gather evidence for the *actus reus* inquiry. For instance, improved versions of *lie detection* techniques – perhaps refined far descendants of the current No Lie MRI or Cephos Corp technologies – might help us to figure out whose testimony to trust and whose testimony to treat with caution. Alternatively, rather than asking people to tell us what happened and then using technology to sieve out the lies, perhaps witnesses might be *compelled* to tell the truth (or at least what they *take* to be the truth) through hypnosis, through the use of refined truth drugs, through the use of techniques like the one being developed by F Andrew Kozel and Mark George that uses transcranial magnetic stimulation to temporarily disable those parts of the brain that are used in intentional deception, or maybe even mind-reading techniques might one day be developed that extract information directly from people's brains. *Brain finger-printing*, like the MERMER-based technique

¹² See N. 10, above, for an explanation of the significance of this qualification.

¹³ The “non-culpably” clause is needed because (as I hinted in N. 10, above) the exculpatory value of incapacities for which we are responsible is usually discounted.

¹⁴ I discuss this in detail in N. 5, above, §3.

(Memory and Encoding Related Multifaceted Electroencephalographic Response), might also help us to identify key witnesses who were present at the scene of the crime, and who might therefore need to be questioned more thoroughly. And finally, perhaps some techniques might also be developed to determine whether a particular person's behaviour qualifies as *genuine actions* — for instance, by uncovering that they can't reliably form- or act upon intentions.

However, neuroscience might also be useful in the *mens rea* inquiry. As I suggested above, among other things responsibility tracks capacity. But since some of our capacities are mental, to the extent that our mental capacities are implemented within our brains,¹⁵ a brain scan which conclusively shows that I lack an important mental capacity – i.e. one that is presupposed by the claim about how I ought to have acted – could therefore at least partially reduce my culpability.¹⁶ This account of how neuroscience might one day fit into the *mens rea* inquiry might be captured by adding just one element to the left hand side of the previous schematic representation:

brain scan → mental capacity → can & can't → duties → culpability

From right to left this schematic reads: one is culpable when one violates one's duties; one's duties depend at least in part on what one can and can't do; what one can and can't do depends at least in part on one's mental capacities; and brain scans might one day reveal people's mental capacities. Thus, for instance, if a brain scan conclusively showed that I lack the capacity for rational deliberation, and this mental capacity was thought to be necessary for moral agency, then that might excuse me for acting culpably. Likewise, if a brain scan showed that I lack another important agency-relevant mental capacity – for instance, an affective rather than a cognitive one, perhaps like Conrad's inability to empathize – then that too might at least partially excuse me. Thus, if a neuroimaging scan of Conrad's brain conclusively showed that he lacks the mental capacity to empathize, and if that capacity was thought to be necessary (e.g.) for someone to truly understand that it is wrong to kill others, then Conrad's reduced capacity for empathy might at least partially excuse

¹⁵ Proponents of the *extended mind hypothesis* and *substance dualists* might object to this.

¹⁶ N. 1, above.

him.¹⁷ After all, we would now surmise that the reason why Conrad killed is because he didn't truly understand that killing is wrong, and people who do not truly understand the reasons that there are to not do certain things can not be expected in the strongest sense of this term to refrain from doing those things, and so they can't be fully blamed for doing those things.

1.4. SUMMARY

Nothing that I said above should be taken as recommending that any of these (and perhaps other) techniques and technologies are ready for "prime time" in courtrooms. In fact, as I have argued elsewhere,¹⁸ there are many scientific, technological, procedural and moral reasons to consider before introducing this or similar technology into courtrooms. However, these considerations do not affect my present rather minimal claim that there *are* ways in which a future neuroscience might help us to assess people's responsibility. Firstly, in much the same way as testimonial evidence and forensic techniques currently help us to answer the "Who dunnit?" question, so too neuroscience may one day provide further pieces of the puzzle for the *actus reus* inquiry. Secondly, just as psychology today provides the law with information about people's mental capacities, so too one day neuroscience may also provide information about people's mental capacities which will help to inform legal determinations of culpability in the *mens rea* inquiry.

2. TWO CONCERNS

In the previous section I explained how a future neuroscience might one day play a role in helping the criminal law to assess people's responsibility; on my account it might do this by providing pieces of the puzzle that inform the *actus reus* and *mens rea* inquiries. An important detail of my account of how neuroscience might fit into the *mens rea* inquiry was the capacitarian idea that responsibility tracks capacity — i.e. that if someone has significant deficits in certain mental capacities, then to the extent that possession of those mental capacities is presupposed by claims about what that person ought to do or not do, the degree of their culpability for what they actually do should be reduced to take account of those deficits.

¹⁷ This would be an instance of extending the application of the *cognitive* requirement of *Daniel M'Naghten's Case*, 8 ER 718, [1843] UKHL J16 to include *affective* incapacities.

¹⁸ I discuss this in detail in N. 5, above, §4.

2.1. HEIDI MAIBOM

However, Heidi Maibom has argued for a picture that looks radically different to the capacitarian one that I endorsed above. On her account, people like Conrad suffer from a “*moral disorder* [in the sense that] in a world with no moral and no social values, [they] would present with relatively few problems. To put it colloquially, [a person like Conrad] is *more bad than mad* [and so] the question of whether [people like Conrad] should be regarded as legally responsible boils down to whether we should excuse those whose mental disorder primarily consists in them *being bad*. That, I hope, is counterintuitive”.¹⁹ Without repeating the details of Maibom’s argument, her account might be schematically represented like this:

brain scan → mental capacity → mad or bad → culpability

From left to right this schematic reads: brain scans may one day reveal people’s mental capacities; once we know their mental capacities we will be able to tell whether they are mad or bad; and once we know whether they are mad or bad we will then be able to tell whether to excuse or to condemn them (i.e. excuse the mad, but condemn the bad). Given this picture, on Maibom’s account the scan of Conrad’s brain would show that he has no room for feelings of empathy – that this is a person who is clearly bad all the way down to his core²⁰ – and if anything people like that should be scorned and condemned rather than excused! Put another way, if Conrad’s history of callous and mean behaviour had not already convinced us that he is a bad man, then the scans of his brain which show that he lacks even the basic capacity to feel empathy for others should cement in our minds the fact that he is a bad man, and people like that should be condemned rather than excused.

Contrary to what I said above, Maibom believes that such evidence of reduced capacity condemns rather than excuses.

2.2. MARGA REIMER

Maibom’s comments may leave the reader perplexed, wondering which of the two interpretations is right — i.e. does such evidence condemn (i.e. what I said) or does it excuse (i.e. what Maibom said). Unfortunately, what Marga Reimer says on the topic

¹⁹ N. 3, above.

²⁰ N. 3, above, 177.

offers little relief from this feeling of confusion, because what she thinks is that such evidence would do both of these things – i.e. that it would *both* condemn *and* excuse – and so she thinks that whether such evidence does one thing or the other will ultimately depend on the language that we use to characterize Conrad’s mental economy (i.e. what capacities he has, to what degree he has them, and what capacities he lacks).

On Reimer’s account, if we characterize Conrad’s situation as involving a *deficit*, disorder or illness of some sort – e.g. a *lack* or a *deficiency* in his affective capacities – then we will tend to view such evidence as a reason to excuse him.²¹ After all, Conrad will be seen as a victim of a mental disease that expresses itself in his killing other people — “It’s not him,... it’s *his disease* that makes him do those terrible things” we might explain. But if we characterize Conrad’s situation using different terminology – for instance, if we say that his emotional coldness is an evolutionary adaptation which allows a few non-cooperators like him to exist in a society populated otherwise by cooperators – then we may instead view such evidence as a reason to condemn him as a self-centred, callous, free-rider who only looks out for number one.²² However, the really troubling thing about what Reimer says is that she doesn’t seem to think that there is any objective fact about whether one characterization is better than the other, or at least this is how I interpret her suggestion that whether we see such evidence as grounds for excuse or as grounds for condemnation will hinge “in effect, [on] a *gestalt shift*. The image formerly seen and characterized as a duck, will now be seen and characterized as a rabbit”.²³ This interpretation of her position is further reinforced by her claim that “*both* defence and prosecuting attorneys might make use of the relevant brain images. Indeed, they might appeal to *the very same images*, but *characterize* them in radically different ways — brains that are ‘damaged’ vs. brains that are ‘evil’”.²⁴

Contrary to what I said above, and indeed contrary to what Maibom said too, Reimer thinks that this sort of evidence *both* excuses *and* condemns at the same time (though not necessarily to the same degree).

²¹ N. 4, above, 185.

²² N. 4, above.

²³ N. 4, above, 190, emphasis added.

²⁴ N. 4, above, 196, emphasis added.

2.3. THE SIGNIFICANCE OF WHAT MAIBOM AND REIMER SAY

On the capacitarian account responsibility tracks capacity, and in §1 I argued that since neuroscience might one day help us to assess people's mental capacities, it might therefore help us to assess their responsibility. However, if what Reimer and Maibom say is right, then the degree of a person's responsibility will not necessarily track the degree to which they possess certain mental capacities – on Maibom's account sometimes a reduction in capacity will result in an *increase rather than a decrease* in culpability; and on Reimer's account some reductions in capacity will *both reduce and increase* a person's culpability – and so measurements of mental capacities will not necessarily translate into the results which I suggested for their responsibility.

3. MY RESPONSE

What Maibom and Reimer say seems to pose a challenge to capacitarianism, and this in turn poses a challenge to what I said earlier about how neuroscience might one day help the law to assess responsibility. However, this section will explain why what Maibom and Reimer say does not in fact challenge capacitarianism. My discussion will begin by drawing attention to two features of their analyses which I find unhelpful; I will then introduce two conceptual tools; and finally those tools will then be employed to defend my account of how neuroscience might one day help the law to assess people's responsibility, and to explain where Reimer and Maibom seem to have taken the wrong turn in their analyses.

3.1. TWO UNHELPFUL FEATURES OF REIMER'S AND MAIBOM'S ANALYSES

A feature of Reimer's analysis which I find unhelpful is her tacit assumption that the things which she says pose a *particular* problem for a neuroscientific approach to responsibility assessment, because I believe that the sorts of worries which she raises are in fact neither new nor specific to the neuroscientific approach.

For instance, reflections about what might be said of someone like Susan Wolf's fictitious JoJo character, might also lead us to be torn about whether to call JoJo "mad" or to call him "bad". To recap, "JoJo is the favorite son of Jo the First, an evil and sadistic dictator [who] is given a special education and is allowed to accompany his father and observe his daily routine[, i]n light of [which] little JoJo takes his father

as a role model and develops values very much like Dad's"²⁵ — i.e. JoJo becomes an evil and sadistic dictator just like his dad. In that paper as well as in a later one²⁶ Wolf notes that one can indeed be torn between two apparently competing assessments of JoJo — on the one hand one wants to condemn him, but on the other hand one can't help but to at least partially excuse him on account of his terribly unfortunate upbringing (after all, few people could say with a clear conscience that if they'd been brought up as JoJo had then they would not have turned out just like him). But the thing to notice here is that we do not need to be presented with any scientific or *neuroscientific* evidence of what JoJo's brain looks like to feel torn about whether to call him "bad" and condemn him, or to call him "mad" and to excuse him (even if only partially), because descriptions of his behaviour and stories about his upbringing will suffice perfectly well to leave us perplexed.

A similar dilemma is also encountered in Gary Watson's reflections about what we might say about someone like Robert Alton Harris, the man who was executed in San Quentin's gas chamber on April 21, 1992, after spending nearly 14 years on Death Row for the cold-blooded murder of John Mayeski and Michael Baker on July 5, 1978. While the brutality and callousness of Harris' actions and his patent lack of any recognisably human feeling or remorse incline one to condemn the man, the tragic story of how a sensitive little Robert was moulded into the hard, mean and nasty Harris by denials of love and experiences of brutality even from his own mother, inclines one to at least partially excuse the man.²⁷

My point is that in both cases — i.e. JoJo and Harris — good old fashioned behavioural evidence and a story about their upbringing will quite effectively leave us feeling perplexed about whether to condemn or to excuse them, and the significance of this is that if there is indeed a problem of the sort that Reimer identifies, then this

²⁵ Susan Wolf, 'Sanity and the metaphysics of responsibility' in (eds.) Ferdinand Shoeman, *Responsibility, Character and the Emotions: New Essays in Moral Psychology* (New York, NY, 1987), 46-62 (page cited 53).

²⁶ Susan Wolf, 'The Reason View' in (eds.) Laura Waddell Ekstrom, *Agency and Responsibility: Essays on the Metaphysics of Freedom* (Boulder, Colorado, USA, 2001), 205-26.

²⁷ Gary Watson, 'Responsibility and the limits of evil: variations on a Strawsonian theme' in (eds.) Ferdinand Schoeman, *Responsibility, Character, and the Emotions* (Cambridge, UK, 1987), 256-86 (pages cited 131-7).

will not just be a problem for the high-tech neuroscientific approach to responsibility assessments that I endorsed above, but it will rather be a problem that affects the good-old-fashioned behavioural approach to responsibility assessment as well. Exactly the same sorts of dilemmas are encountered irrespective of whether we are presented with neuroscientific evidence or just behavioural and historic evidence, and so the fact that neuroimaging evidence might leave us troubled in this way is not a reason to suppose that the high-tech neuroscientific approach to responsibility assessment suffers from a unique flaw that does not afflict other approaches to assessing responsibility as well.

On the other hand, a feature of Maibom's analysis which I find unhelpful is her use of the *mad-bad* dichotomy in the context of discussing responsibility. Throughout her discussion, Maibom revisits the question of whether such neuroscientific evidence would establish that people like Conrad are mad and thus that their responsibility is diminished (i.e. reduced culpability), or whether it would establish that they are bad and thus that their responsibility is not diminished (i.e. maybe even greater culpability). Maibom of course thinks that such evidence would exclusively show that such people are bad, hence her insistence that such evidence would condemn. However, what is implicit in her use of this *mad-bad* dichotomy – and this is the feature to which I now draw attention – is her implicit assumption that once we establish whether a person is mad or whether they are bad, this will then tell us something about that person's responsibility. Put another way, the operative assumption behind Maibom's use of the *mad-bad* dichotomy is that claims about whether someone is mad or bad provide premises in support of conclusions about their culpability. But what troubles me is that on the capacitarian account this operative assumption gets things completely the wrong way around, because what we should do *first* is to determine whether a particular incapacity affects someone's status as a moral agent, and only once we have already determined whether their status as moral agents is compromised might we *then* be in a position to call wrong-doers whose agency is compromised "mad" and to call those whose agency is not compromised "bad".

There are two opposite ways of viewing the grounding relationship between claims about responsibility and claims about madness and badness: on one view claims about madness and badness *ground* claims about responsibility, and on the other view

claims about madness and badness are *grounded by* claims about responsibility. Evidently, Maibom thinks that the former view of this grounding relationship is correct. However, given that capacitarions would endorse the latter view, it is therefore highly inappropriate for Maibom to at the outset adopt a hostile analysis of the grounding relationship between claims about responsibility and claims about madness and badness, since by doing this her argument effectively begs the question against capacitarions. Put another way, my point is not that the view which Maibom adopts is wrong, or that the view which capacitarions adopt is right, but it is rather that if adopting either view would unfairly bias our analysis either side then an altogether different way of discussing responsibility should be adopted.

This objection to Maibom's use of the *mad-bad* dichotomy in her analysis has a counterpart in the sorts of problems that are encountered when we think through similar issues by using the *disease-person* dichotomy. As Robert Sapolsky notes, when we consider "individuals who have committed grotesquely violent, sociopathic crimes, but who have no demonstrable [brain] damage...", it seems a fatuous tautology to say that there *must* be an organic abnormality in such cases – 'it is *only* an organically abnormal brain that produces [such patently] abnormal behaviour' – and that we simply lack sufficiently sensitive techniques for demonstrating it".²⁸ Sapolsky's comments suggest that when we are confronted with people like this, it is at least as easy to view their vile actions as symptoms of some underlying though presently undetectable organic disease – the thought here would be something like "No *healthy* person could be like that, right?" – as it is to view those actions as expressions of their evil personality. However, as the recent debate in response to a paper by John Z. Sadler in the journal *Philosophy, Psychiatry & Psychology* highlights,²⁹ it is very difficult to classify people's behaviour in a value-neutral manner as either a symptom of disease/damage/disorder or as an expression of their vice, evil or perversity of character. As Sadler notes in a concluding collective response to the various authors who commented on his paper, "[e]very misbehavior

²⁸ Robert M. Sapolsky, 'The frontal cortex and the criminal justice system' (2004) 359 *Philosophical Transactions of the Royal Society of London* 1787-96 (page cited 1793, emphasis added)

²⁹ John Z. Sadler, 'Vice and the Diagnostic Classification of Mental Disorders: A Philosophical Case Conference' (2008) 15 *Philosophy, Psychiatry & Psychology* 1-17.

has its psychology, and if we proclaim [it] abnormal or deviant, *then* we have a ... dysfunction”³⁰ — i.e. it is only in virtue of the pre-judgment that the behaviour in question is abnormal or deviant that we can even arrive at the subsequent judgment that what we have in front of us is a dysfunction. Potter and Zachar put this point by saying that although psychologists and psychiatrists should indeed take a serious stab at trying to specify the “underlying pathological processes” of disorders, they also recognize that the problem with attempting to do this is going to be that “defining impairment in terms of underlying processes *requires having some conception* of what counts as normal”³¹ — i.e. yet again we are urged that an evaluative exercise is the precursor to deciding whether something is a disease or a disorder. Finally, as Stephen Morse puts it, “values always play a role in the definition of mental disorder”.³² My point is that whether we classify a particular behaviour in one way (as a disorder, disease, damage or illness) or in another (as an expression of that person’s character) is not a straight-forward matter – it is *not* something that might, for instance, be discovered under a microscope by a pathologist in a lab, or during a health check-up with your physician – because the physical facts under-determine our characterization.

Admittedly, there are some cases in which the *disease* interpretation suggests itself more naturally — for instance, the cases of Phineas Gage and of Burns’ and Swerdlow’s³³ 40 year old patient who suddenly developed paedophilic urges due to a brain tumour that pressed on the orbitofrontal region of his brain immediately come to mind. Both of these examples involve trauma and sudden changes which we intuitively view as “damage” and as “disease”, and so it seems more appropriate to attribute the observed changes in such people’s behaviour to the *disease* rather than to these *people*. The people, we might suppose, lie buried somewhere deep beneath the

³⁰ John Z. Sadler, ‘Cause, Fault, Norm’ (2008) 15 *Philosophy, Psychiatry & Psychology* 51-5 (page cited 52, emphasis added).

³¹ Nancy Nyquist Potter and Peter Zachar, ‘Vice, Mental Disorder, and the Role of Underlying Pathological Processes’ (2008) 15 *Philosophy, Psychiatry & Psychology* 27-9 (page cited 28, emphasis added)

³² Stephen J. Morse, ‘Vice, Disorder, Conduct, and Culpability’ (2008) 15 *Philosophy, Psychiatry & Psychology* 47-9 (page cited 48).

³³ Jeffrey M. Burns and Russel H. Swerdlow, ‘Right Orbitofrontal Tumor With Pedophilia Symptom and Constructional Apraxia Sign’ (2003) 60 *Archives of Neurology* 437-40.

diseases, heavy under their weight and unable to move – or maybe the real people no longer even exist because they were destroyed by the damage or by the disease – which is why it seems inappropriate to attribute such behavioural changes to those people. However, what should we say about cases in which *congenital* brain differences result in similar behaviour to that exhibited by Burns and Swerdlow’s patient? Should the behaviour of such “natural born paedophiles” – i.e. people who have *always* had brains that function in a similar way to Burns’ and Swerdlow’s patient’s brain when it was compressed by the tumour – be attributed to a disease even though their brains may have always been like that, or should their behaviour in such cases be attributed to them? Alternatively, how about people who *develop* certain brain features *in early life* because of their terrible upbringing – perhaps people like Robert Alton Harris and JoJo – which predispose them later in life to committing atrocities; is their behaviour due to a disease of their mind, or is it an expression of their evil selves?

In these latter two cases we are in danger of falling prey to the gestalt switch that Reimer describes, one moment viewing these people’s actions as expressions of their rotten character, and the next viewing them as expressions of a disease of the mind (albeit a disease of the mind which they have always had). Lawrence Tancredi’s solution to this problem is to suggest that since the brains of people like Harris and JoJo may in the end be little different to the brains of people like Phineas Gage and of Burns’ and Swerdlow’s patient, if in the former cases we would attribute the behaviour to damage and disease, then for the sake of consistency we should do likewise in the latter cases too.³⁴ However, the arbitrariness of the way in which Tancredi chooses which side is the baseline worries me — perhaps instead of taking the disease interpretation as the baseline and then claiming that all bad people should be re-classified as mad, maybe instead we should take the person interpretation as the baseline and then claim that all mad people should be re-classified as bad. My aim is not to seriously suggest that we should indeed embrace the latter suggestion, but it is rather to point out that sheer *consistency* only requires us to jump one way, but it does not yet tell us which way to jump, and so for this reason I do not find appeal to the *disease–person* dichotomy any more helpful than the *mad–bad* dichotomy. Both

³⁴ Lawrence R. Tancredi, ‘The Bad and the Mad’ in *Hardwired Behavior: what neuroscience reveals about morality* (New York, NY, 2005), 143-61.

dichotomies presuppose contestable substantive decisions about the accused person's status as a moral agent, and hence both of these dichotomies lack the sort of impartiality that is required to settle disputes about responsibility.

To sum up my disenchantments, the worries which Reimer raises are general worries about all responsibility assessments, and hence they do not pose any specific new challenge to the high-tech approach which I endorsed above. Secondly, my objection to Maibom is that the *mad-bad* and *disease-person* dichotomies presuppose certain judgments about people's status as moral agents, and hence such dichotomies should not be used as premises in support of conclusions about responsibility since doing so begs the question against capacitarrians.

3.2. SOME CONCEPTUAL TOOLS: TWO KINDS OF CULPABILITY, AND TWO STAGES OF A CRIMINAL TRIAL

I now introduce the two conceptual tools which will be used to defend my account of how neuroscience might one day help the law to make responsibility assessments.

The first conceptual tool is a distinction between two responsibility concepts – what I call “virtue responsibility” and “capacity responsibility” – which relate to culpability for two different kinds of things. To see this distinction, consider the following example about my two fictitious children, Jane and John:

Jane is 8 years old and very well behaved. She gets up in the morning all by herself, she washes and gets dressed and even makes her own lunch, she doesn't fight with other kids at school, after school she does her homework, she cleans up after herself, she helps me make dinner, and she even looks out for her older brother John. However, despite the fact that Jane is such a responsible little girl, I also know that she is ultimately only a little girl – she still lacks the mental capacities that one needs to have in order to be a fully responsible person – and so I do not really blame her when she sometimes fails to do these things. Jane is a responsible person, even though she is not yet a responsible person; or put in a less ambiguous way, Jane is a responsible little girl, even though she is not yet a fully responsible person.

Now consider John, Jane's older brother. John is 17 years old and not at all like his much younger sister – some would call John an irresponsible young man. He wags school, he won't make his own lunch even though bought lunch is much more expensive, if and when he gets to school he always gets in some kind of trouble with the teachers, his bedroom is a pigsty and we are always cleaning up after him in the

rest of the house; he can not even be trusted to look after his little sister. Nevertheless, despite the fact that John is an irresponsible young man, we all know him well enough to know that he is actually a fully responsible person – we’ve seen what he is capable of doing when he puts his mind to things – and that is precisely why we are so very dark on him for his constant misbehaviour. John is not a responsible person, even though he is a responsible person; or put in a less ambiguous way, John is an irresponsible young man, even though he is in fact a fully responsible person.³⁵

The point of this example is to demonstrate that we can make two similar-sounding but ultimately very different kinds of responsibility assessments. Sometimes, we make assessments about what a person is like – i.e. whether they are good or bad; culpable, admirable or just normal like most other people; or what I call their *virtue responsibility* – and it is in this virtue responsibility sense that Jane is said to be a responsible little girl while John is said to be an irresponsible young man. But at other times we make assessments about what I call *capacity responsibility* – these are assessments about whether the person in question has the mental capacities that it takes to even be a moral agent or a legitimate candidate for attributions of blame for the things that they do – and it is because of our different judgments about their capacity responsibility that Jane is said to not be fully responsible for what she does whereas John is said to be responsible for what he does. This distinction is not particularly novel. In his paper *Two Faces of Responsibility* Gary Watson also argues that we should distinguish these two kinds of responsibility assessments from one another, except that he gives them different names: what I call “virtue responsibility” he calls the *aretaic* face of responsibility, and what I call “capacity responsibility” he calls the *accountability* face of responsibility.³⁶ But putting aside terminological differences, my first point is simply that if we accept the legitimacy of this distinction, then we will notice that people can be culpable on account of two very different kinds of things. Firstly, people can be culpable on account of *what they did*, but secondly, they can also be culpable on account of *who they are*, and as the Jane and John example demonstrates, saying that someone is culpable in one sense does not yet

³⁵ Nicole A Vincent, ‘Responsibility: distinguishing virtue from capacity’ (2009) 3 *Polish Journal of Philosophy* 111-26 (pages cited 115-6).

³⁶ Gary Watson, ‘Two Faces of Responsibility’ in *Agency and Answerability* (Oxford, UK, 2004), 260-88.

commit us to saying anything about whether they are culpable in the other sense since these two judgments can come apart.³⁷

The second conceptual tool that I now introduce is the observation that these different assessments of responsibility and culpability play a role at different stages in the criminal trial. At the *guilt determination* stage, where the focus is on determining whether the accused party is responsible for some outcome or state of affairs – i.e. where the focus is on assessing the sort of responsibility that was the topic of discussion in §1 above – what matters is the degree of the accused party’s *capacity responsibility*. After all, the whole point of the insanity defence (and the other psychological defences that were cited in the main text of that section) is to establish that the party in question does not satisfy all of the criteria of a fully-fledged moral agent, in virtue of which their responsibility for the said outcome or state of affairs should be reduced. However, once we progress to the *sentencing* stage, as we ponder how that person should now be treated on account of what they have done and on account of the degree of their responsibility for it, we may take into account mitigating and aggravating factors, and one such factor is precisely whether that person is responsible or irresponsible in the *virtue responsibility* sense. After all, we may choose to be lenient on a person who has always been responsible in the virtue responsibility sense and for whom this was their first transgression, but a repeat offender who has a history of irresponsibility (yet again in the virtue responsibility sense of the term) may be given a harsher sentence. Thus, while assessments of people’s *capacity responsibility* and the related culpability for *what they did* play a role during the *guilt determination* stage, at the *sentencing* stage those things are taken for granted and our focus shifts to assessments of people’s *virtue responsibility* and to culpability for *who they are*.

With these two conceptual tools in place – i.e. the distinction between the two kinds of culpability assessments, and the different stages of the criminal trial at which each assessment plays a role – I will now defend my earlier claims about how a future neuroscience might help the law to assess responsibility, and I will also comment on where I think Reimer’s and Maibom’s analyses went wrong.

³⁷ N. 35, above, 124-5.

3.3. A DEFENCE OF MY POSITION

I said earlier that Maibom's and Reimer's concerns seem to cast doubt on the capacitarian claim that responsibility tracks capacity, and that with the capacitarian idea shaken up, it no longer seems to make sense for me to suggest that responsibility might one day be assessed in part by using neuroscience to measure people's mental capacities. However, I believe that when the two points made immediately above are taken into account, we will see that Maibom's and Reimer's concerns do not in fact cast doubt on the capacitarian claim that responsibility tracks capacity.

Maibom's comments were meant to cast doubt on the capacitarian claim that responsibility tracks capacity by showing that sometimes (e.g. in Conrad's case) a reduction in capacity would result in an increase rather than a decrease in culpability. The idea here was meant to be that if sometimes a reduction in capacity will result in an increase rather than a decrease in culpability, then sometimes responsibility would *track* capacity but on other occasions responsibility would *inversely track* capacity. However, on closer inspection we should notice that on the *reverse tracking* occasions a different sense of responsibility is in play to the sense which is in play on the *normal tracking* occasions. Specifically, when Maibom insists that Conrad's reduced capacity to empathize would be a reason to condemn him, what she means is that Conrad is culpable on account of *who he is* — i.e. the *virtue responsibility* sense is in play here. However, in asking whether Conrad's reduced capacity to empathize should excuse him at the guilt determination stage, our interest is in his culpability for *what he does* — i.e. our interest is in his *capacity responsibility* not in his virtue responsibility — and so the fact that his culpability on account of *who he is* goes up rather than down is hardly a counter-example to the capacitarian claim that his responsibility for *what he did* would go down. Put another way, Maibom might indeed be right that on some occasions a reduction in capacity will result in an increase rather than a decrease in culpability; however the increase will be in culpability for *who that person is* (at the sentencing stage) and not for *what that person did* (at the guilt determination stage), and so this is not a counter-example to the capacitarian claim that responsibility *for what a person does* tracks capacity.

Turning now to Reimer, her comments were meant to cast doubt on the capacitarian claim that responsibility tracks capacity by showing that some reductions in capacity will both reduce and increase a person's culpability, rather than only

reducing their culpability. But, using Conrad as the example, once we notice that we would excuse him *for what he did* whereas we would condemn him *for who he is* – as well as the fact that the excuse would occur at the guilt determination stage, whereas the condemnation would show up at the sentencing stage – there is again little reason to view Reimer’s observation that *both* excuse *and* condemnation are on order as a counter-example to the capacitarian claim that responsibility tracks capacity, since the claim that responsibility tracks capacity only applies to responsibility for what a person does — i.e. it only applies to our assessments of responsibility at the guilt determination stage. Put a different way, the sense in which Conrad’s responsibility would be increased (i.e. we would condemn him on account of *who he is*) is completely different to the sense in which his responsibility would be decreased (i.e. we would excuse him, even if only partially, for *what he did*). Furthermore, the increase would happen at a different time (i.e. during sentencing) than the decrease (i.e. during guilt determination), and so there is no need (and in fact no possibility even) to offset the excuse against the condemnation.

Thus, in addition to the problems which I discussed in §3.2. above, Reimer’s and Maibom’s analyses also seem to have gone wrong in two further ways. Firstly, they both failed to distinguish the two very different sorts of things on account of which someone might be judged culpable — i.e. that people can be culpable on account of *what they do* or on account of *who they are*; and secondly, they both failed to give due recognition to the fact that while the excuse would play a role at the *guilt determination* stage, the condemnation would play a role at the *sentencing* stage of the criminal trial. Taken together, these two problems lead to the following three shortcomings: firstly, both of them failed to notice that there is *no genuine conflict* between the excuse- and the condemnation interpretation of the neuroscientific evidence;³⁸ secondly, this oversight led Maibom to think that she must choose between *either* saying that psychopaths should be condemned *or* saying that they should be excused (and she chose the former); and thirdly, this oversight also led Reimer to think that the exculpatory effects of reduced capacities *would be reduced* by their condemnatory effects.

³⁸ Putting the same point, though in the terminology of the *mad–bad* dichotomy, Morse also insists that “[a]ctions can be *just* mad, *just* bad, or mad *and* bad” N. 32, above, 47, emphasis added.

I now conclude that neither Maibom's nor Reimer's comments pose a problem for the capacitarian claim that responsibility tracks capacity, and I see this as a defence of my earlier account of how neuroscience might one day play a role in helping the law to assess people's responsibility.

4. CONCLUSION

This paper has advanced two points. Firstly, I argued that although Reimer is right in claiming that some evidence of reduced capacity *both* condemns and excuses, she is wrong in thinking that such evidence condemns people *for the same thing* as the thing for which it excuses them. On my account, what such evidence *condemns* people for is *for who they are*, whereas it *excuses* them *for what they do*. This observation is important because it helps us to notice that such evidence would support the *defence* at the *guilt determination* stage (i.e. when our interest is in assessing the degree of a person's responsibility for what they have done), whereas it would support the *prosecution* only at the *sentencing* stage (i.e. when we are interested in assessments of their character as either a mitigating or an aggravating factor).

Secondly, I also rejected Maibom's claim that some evidence of reduced capacity condemns *rather than* excuses. Like Reimer, Maibom too fails to notice that such evidence condemns people *for different things* to the things for which it excuses them, and that it does this *at different stages* of the criminal process. But unlike Reimer, who was driven by these oversights to conclude that both of these assessments must therefore be valid, Maibom instead *took sides* and insisted that only one of them is right and that the other is mistaken. However, I have argued that there is *no need to choose* between *either* condemning *or* excusing people with incapacities like Conrad's, because while the condemnation pertains to their *who they are* and is relevant to the *sentencing* stage, the *exculpation* pertains to *what they did* and applies to the *guilt determination* stage of a criminal trial.

A positive consequence of these two largely negative arguments is that at the guilt determination stage a finding which conclusively shows that Conrad lacks such a capacity may indeed reduce the degree of his responsibility for what he did – i.e. such evidence would not testify for both the defence and for the prosecution at the guilt determination stage, but only for the defence – however admittedly at the sentencing stage such evidence may indeed condemn him on account of who he is.