***Deep Responsibility and “Morality”[[1]](#footnote-1)***

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Paul Russell

Lund University

July 2024

[Forthcoming in Michael Frauchiger & Markus Stephanians, eds., *Themes from Susan Wolf*. De Gruyter.]

*A satisfactory theory of (deep) responsibility must not only be able to identify which agents are responsible, and for what - it must be able to explain why they are responsible, and, ultimately, why the idea of responsibility makes any sense at all.*

- Susan Wolf[[2]](#footnote-2)

Among Susan Wolf’s most distinguished and influential contributions to philosophy has been her work concerning problems of free will and moral responsibility. Much of this work was done early on in her career in the form of several substantial papers that culminated in her 1990 book *Freedom Within Reason* (hereafter abbreviated as FR).[[3]](#footnote-3) In *Freedom Within Reason* Wolf develops a distinct and illuminating position on this subject – a position that she refers to as “The Reason View”. This account of free will and moral responsibility is constructed around an interpretation of the relationship between responsibility and rationality and presented in support of compatibilism. Wolf’s Reason View is one of several such compatibilist accounts presented in last two decades of the 20th century.[[4]](#footnote-4) Taken together, these accounts serve as an important movement within the contemporary free will debate, providing a more sophisticated alternative to earlier compatibilist views.

What particularly distinguishes Wolf’s Reason View from other compatibilist positions that emphasize the relevance of rational competence for moral responsibility is Wolf’s defence of an “asymmetrical” account of moral responsibility.[[5]](#footnote-5) The asymmetry involved concerns the claim that conditions required for praise and for blame diverge with respect to the requirement that the agent had the ability to do otherwise. While this ability is required for blame, it is not, Wolf maintains, required for praise (FWR,68-9, 79, 89). On the face of it, this feature of the Reason View puts some strain on Wolf’s fundamental aim, which is to vindicate our ordinary or “mundane” understanding of the concept of moral responsibility and to defeat the sceptical threat that it is somehow “incoherent” or impossible to satisfy (FWR, vi, 3, 6, 21-22, 26, 132, 144-7; Wolf, SMR, 281-2).[[6]](#footnote-6) The particular concept of moral responsibility that Wolf aims to vindicate is what she refers to as “deep responsibility” (FWR, 10-11, 40-5, 63-4). According to Wolf, when “we consider an individual worthy of blame or praise, we are not merely judging the moral quality of the event with which the individual is intimately associated; we are judging the moral quality of the individual herself in some more focused, noninstrumental, and seemingly serious way.” (FWR, 41; and also 19, 34). Individuals who we take to be deeply responsible are understood to be *accountable*, where this raises the question of whether an individual *deserves* praise or blame (FWR, 19-20, 43-4, 63-4, 66). [[7]](#footnote-7)

***I. Real Selves and Deep Selves***

How, then, are we to make sense of this conception of “deep responsibility”? Wolf’s approach to this question involves a critical and a constructive aspect. The critical aspect takes the form of a consideration of two views that she rejects – both of which she regards as insufficiently connected to our everyday, “prephilosophical” understanding of the conditions relating to free will and responsibility (SMR, 281-2). The first to be considered is a compatibilist proposal that aims to improve on classical compatibilist views that reduced free will and responsibility to accounts of freedom of action or doing what one wills. Clearly, we need more than an ability to control our actions by means of our desires and choices (FWR, 27-8; SMR, 283). We also want freedom of will as well as freedom of action, where this is secured by means of having some sort of control over those desires that shape and direct our will and lead to action.

There are several compatibilist accounts of freedom of will that fall under the heading of what Wolf labels the “Real Self View”.[[8]](#footnote-8) Responsible agents, unlike animals, children, and others like them, are able to govern their wills in light of their “valuational system”, where a person’s values are “what that person thinks good or supported by reason” (FWR, 31). In these circumstances an agent’s actions are governed by her “real self”. When, according to this view, we can attribute an agent’s behaviour to her real self this serves as a necessary and sufficient condition of responsibility (FWR, 34). Nevertheless, this theory faces a significant problem: who, or what, is responsible for the agent’s real self? (FWR, 37; and SMR, 285) The Real Self View, Wolf argues, “seems only to have pushed the problem further back” (SMR, 285). Agents that satisfy these conditions still seem subject to fate and inevitability, and the addition of “another loop to the internal structure of the agent” cannot provide the sort of control that we are looking for here (FWR, 44).

Wolf illustrates what is still missing from the Real Self View, with a vivid - and influential - example. In “Sanity and the metaphysics of responsibility” she describes the case of “JoJo” (SMR, 288-93; and compare the parallel example of “Tony” at FWR, 112-3). JoJo is a person who has been raised in such a way that he acquires his moral values from his father, who is brutal and cruel dictator. As an adult, JoJo behaves in much the same cruel and brutal way that his father did before him. Due to his upbringing JoJo has a “real self” that he has no ability to change or “correct”. For this reason, it is not possible for JoJo to “know the difference between right and wrong”, nor can he “revise” the (faulty) values that he has been indoctrinated with by his father and those around him. Even though JoJo may satisfy the conditions of responsibility proposed by the Real Self View it is evident, according to Wolf, that JoJo is *not* responsible. Cases like JoJo, Wolf argues, are not only relevantly similar to cases of childhood deprivation, they also serve explain “why we give less than full responsibility to persons who, though acting badly, act in ways that are strongly encouraged by their societies – the slaveowners of the 1850s, the Nazis of the 1930’s, and many male chauvinists of our father’s generation, for example” (SMR, 290).[[9]](#footnote-9) All these agents, although they may have a Real Self, are still (intuitively) not responsible. The kind of responsibility they have, Wolf argues, remains “superficial” (FWR, 40-1, 80-1).[[10]](#footnote-10) So what is still missing?

Before offering her Reason View account, Wolf first examines an alternative proposal, the “Autonomy View”. According to this view, no adequate conception of responsibility can leave us subject to the arbitrary and accidental conditions of heredity and environment (FWR, 12-4, 47, 76-7). Deeply responsible agents are not “the inevitable products” of “the blind ceaseless flow of the world’s events” but rather possess a power to “transcend” their existing values and become the *ultimate* source of their values and actions (FWR, 10-15, 44, 50-5). The problem with this account, Wolf argues, is that a “radical freedom” of the kind that the Autonomy View takes to be required to justify deep responsibility is incoherent. It is, moreover, something that “no one could ever have reason to want to exercise” (FWR, 55). “To want autonomy”, Wolf suggests, “is to want not only the ability to act rationally but also the ability to act irrationally – but the latter is a very strange ability to want, if it is an ability at all.” (FWR, 54-5, 56, 62, 66, 68-9, 140). Why should we want to be able to act *against* Reason, as well as with it?[[11]](#footnote-11)

***II. The Reason View, Depth and Asymmetry***

The Reason View suggests a more moderate position that lies between the extremes of the Real Self View and the Autonomy View (FWR, 66, 68, 78). On one side, the conditions offered by the Real Self View are too weak and, on the other, the conditions offered by the Autonomy View are too strong. We need more than to be able to act according to the dictates of our Real Self, and we need less than a “radical freedom” of the sort that the Autonomy View demands. More specifically, we need to be able to account for *depth* without collapsing into incoherent metaphysics. The ability that is crucial to responsibility, Wolf maintains, is simply the ability to act in accordance with Reason.

According to the Reason View, what we need to know if we are to find out whether we are free and responsible beings is whether we possess the ability to act in accordance with Reason. Since Reason is here understood to refer to the highest faculty or set of faculties there is, the faculty or set of faculties that, in most circumstances, will help us form true beliefs and good values, this amounts to the suggestion that we need to know whether we have the ability to think – and on the basis of our thought, to act – well rather than badly. That is, we need to know whether we have the ability to choose and to act on the basis of the right reasons for choosing and acting.” (FWR, 70-1)

This is not so much a matter of having a distinct metaphysical power (“self-creation”, etc.) but of possessing “a distinctive intellectual power, the power to exercise the right Reason and to govern one’s actions accordingly” (FWR, 71) In this way, the freedom and power necessary for responsibility “are the freedom and power to be good, that is the freedom and power to do the right thing for the right reason” (FWR, 77). We do not, therefore, need some (incoherent) kind of control over our Real Selves so that they are somehow (ultimately) “up to us”– we “need simply to *be* a certain way, even though it is not within our power to determine whether we are that way or not” (SMR, 288 – Wolf’s emphasis). What is missing in the case of JoJo and others like him, therefore, is not that he is not the ultimate source of his own Real Self but that he is unable “to see and appreciate the world for what it is” (FWR, 93).

Although Wolf aims to capture our ordinary common sense understanding of moral responsibility, the Reason View is, nevertheless, committed to a claim that Wolf concedes is “curious” (FWR, 79).[[12]](#footnote-12) As we have noted, for the Reason View it is not required that “the agent had the ability to act in *dis*cordance with Reason” (FWR, 68 - Wolf’s emphasis). However, when the agent *fails* to do what she ought to do, then the question of whether she could have done otherwise becomes “all important” (FWR, 68-9). This presents us with an “asymmetry” between the conditions for praise and blame. Contrary to the Autonomy View, praise does not require that the agent could have done otherwise. So long as the agent is able to do the right thing for the right reason they satisfy the requirements of the Reason View, and hence agents who do the right thing for the right reason *deserve* praise.[[13]](#footnote-13) In situations of this kind the agent possesses the ability to act in accordance with the true and the Good, and this suffices for responsible conduct.

Blame involves moral failure, where the action or conduct in question deviates from what Reason requires of us. The question that arises in these circumstances is whether or not it is *always* the case that an agent who does the wrong thing could not do the right thing for the right reason? The case of JoJo suggests that there are *some* agents who act badly or do the wrong thing but do not deserve blame. Due to their upbringing or deprived childhoods these agents are unable to be guided by the True and the Good. Although JoJo and those like him (e.g. Nazis, slave owners, male chauvinists et al), may all have a Real Self, they are still unable to “correct” or “revise” their values in light of the True and the Good (SMR, 293; FWR, 85, 92-3, 112-3, 125). Although they can govern their conduct by the reasons provided by their Real Self, they are not able to “be governed by what reasons there are” (FWR, 93). However, this is not, Wolf argues, true of all “bad-acting agents”.

Two things need to be true of an agent who is able to do what is right. She must be able to *recogniz*e the right thing to do and she must be able to *execute* it (FWR, 80).

... we may characterize the ability to X as consisting of two claims, one positive and one negative. The positive claim is that the individual to whom the ability is attributed possesses whatever capacities, skills, talents, knowledge, and so on are necessary for X-ing. ... The negative claim is that nothing interferes with or prevents the exercise of the relevant capacities, talents, and so on.... In light of this characterization, the claim that A is unable to X can be understood as the claim that either A lacks at least one of the capacities, talents, skills or whatever that are necessary for X-ing or something prevents or interferes with A’s exercise of these capacities, talents, or skills. (FWR, 101)

With this account of ability and inability to X in place, Wolf argues that while *psychological* determinism is incompatible with responsibility for wrong actions, this is not true of *physiological* determinism (FWR, 101-03). The intuition that supports Wolf’s contrast between the implications of psychological and physiological determinism is that the former is akin to cases of responsibility undermining factors of a kind that we find in cases of hypnotism, addiction, and such, while the latter has no such implications (FWR, 100-02, 105, 139).[[14]](#footnote-14)

Wolf offers another example to support her case for asymmetry and the Reason View. “Rose” is a Professor who chooses to watch television instead of grade papers, which is what she ought to do. Rose does the wrong thing. However, unlike JoJo or Tony, Rose not only possesses the general ability to recognize and respond to considerations relating to what she ought to be doing, there is nothing that “interferes with or prevents her from exercising these skills, talents, and capacities” (FWR, 111, 145). She is not, Wolf suggests, psychologically determined to act badly or do the wrong thing. Moreover, even if *physical* determinism were true, Wolf argues, this would not “prevent” or “constrain” her from doing the right thing (FWR, 101-02, 108-09, 114).[[15]](#footnote-15) Wolf contrasts this with the case of Tony, who is *psychologically* prevented by his upbringing and background from being able to do what is right. Unlike Rose, Tony is not psychologically free to choose otherwise and, therefore, he is not fully responsible (FWR, 113). What the example of Rose shows, according to Wolf, is that there are cases of wrongdoing or acting badly where the agent is still able to satisfy the fundamental condition of responsibility. That is to say, Rose, and those like her, retain an ability to act in accordance with the True and the Good and they can, therefore, be “fairly held accountable” or blamed when they do wrong (FWR, 104). In these circumstances the agent is appropriately blamed because, although they acted badly, they could have done otherwise.[[16]](#footnote-16) While it is true that the requirement to be able to do otherwise applies to cases of blame but not to praise, there is still, Wolf argues, “a single set of conditions for responsibility for any type of act”: namely, that the responsible agent has an ability to act in accordance with the True and the Good (FWR, 90).

***III. Reflective Self-Control and “Morality”: Wallace’s Theory***

Wolf’s Reason View, as we noted, belongs to a group of similar views that defend forms of compatibilism that are constructed around the relevance of rational capacities for understanding the nature and conditions of moral responsibility. Among the most influential of these has been R. Jay Wallace’s *Responsibility and the Moral Sentiments*, which was published a few years after Wolf’s *Freedom Within Reason*. Like Wolf, Wallace rejects the Autonomy View or “strong freedom of will”, which emphasizes a “metaphysical interpretation” of what it is to be morally responsible (Wallace, 1994: 85-7). The “picture” of responsibility that Wallace defends invites us to consider when it is *fair* to hold people responsible (Wallace, 1994: 5-6, 15-16, 85, 93-4, 101-09, 114-5, 135, 149, 215, 227). What responsible agents need, he argues, is to possess “powers of reflective self-control”, which is compatible with determinism (Wallace, 1994: 86-7, 182, 194, 220, 232). According to this view, “normative competence” or reflective self-control requires: (1) the power to grasp and apply moral reasons, and (2) the power to control or regulate behaviour in light of such reasons (Wallace, 1994: 86, 157-61, 220-1). In respect of these matters Wolf and Wallace are largely agreed.

There are, nevertheless, some significant differences between Wallace’s “reflective self-control” account and the Reason View’s understanding of what is required to satisfy conditions of responsible agency (FWR, 80, 86-7, 101, 111). There is, in particular, a significant difference between them in respect of the issue of asymmetry and the principle of alternate possibilities. According to Wolf, if “an agent is incapable of doing the right thing for the right reasons, then it is not her fault that she stumbles into doing something wrong” (FWR, 81). Justified blame, Wolf claims, presupposes that the agent concerned could have done otherwise (i.e. had alternate possibilities) (FWR, 85-9, 89-90, 112-4). Although Wolf does not frame it in these terms, what drives this further requirement, is the sense that it would be *unfair* to blame (and punish) agents who lack the opportunity to do otherwise or who cannot avoid blame.[[17]](#footnote-17) Wallace refers to any view of this kind as a “bipartite account” and he rejects it (Wallace, 1994: 208). What matters for moral responsibility, he argues, is not any power to avoid blame or “specific ability to exercise their general powers in the circumstances of action” (Wallace, 1994: 199, 201). All that matters is simply that the responsible agent *possesses* such powers - whether they can *exercise* them differently in the particular circumstances or not (Wallace, 1994: 182-3).

Why, then, does Wallace reject “bipartite accounts” and the principle of alternate possibilities? It is, Wallace argues, “a fatal mistake” to assume the burden of satisfying “avoidability” or include “opportunity” among the conditions of responsibility (Wallace, 1994: 187-8, 196-207). When compatibilists “grant that responsibility requires opportunities or alternative possibilities or alternate possibilities, they may already have given the game away” (Wallace, 1994: 223). In order to satisfy conditions of this kind within compatibilist constraints we would need to distinguish different senses of opportunity or alternate possibility that are, Wallace claims, “too fine-grained and technical to do the normative work required of them” (Wallace, 1994: 223). These conditions are, in any case, unnecessary for understanding the forms of abilities and powers required for moral responsibility. What matters for the purpose of judging if an agent has an ability to do something is not that they can exercise the power or ability in particular circumstances but that they possess the relevant (general) ability or power (Wallace, 1994: 183, 190). We can, Wallace maintains, appreciate this by way of analogy to other cases. It is, for example, unreasonable to expect a new immigrant to speak a language as fluently as a native, or to expect someone who lacks any athletic talent to excel at sports (Wallace, 161-82). But it is entirely reasonable to expect a professional athlete to, say, catch a “very catchable” ball even if they drop it on some particular occasion or play (Wallace, 1994: 192). The same applies for agents who fail to do the right thing on some occasion but who, nevertheless, still possess relevant general powers of reflective self-control. In the absence of excusing considerations, it is still fair to blame them, whether or not in the actual conditions they “could not have done otherwise” or lacked the opportunity to do what is right.[[18]](#footnote-18) In sum, from Wallace’s perspective the Reason View leaves compatibilism hostage to sceptical doubts because it makes demands that are unnecessary, as well as “vulnerable to incompatibilist counterattack” (Wallace, 1994: 193).

Another important feature of Wallace’s reflective self-control model that we need to take note of, in respect of assessing his views in relation to Wolf’s Reason View, is that Wallace makes clear that his understanding of moral responsibility is consistent with what Bernard Williams has described as “the morality system” (Wallace, 1994: 39-40, 64-6). Whereas Williams aims to discredit the morality system, by means of a genealogical critique, Wallace takes his reflective self-control model to provide a vindication of “morality’s” understanding of moral responsibility. To explain this we need, first, to describe how Williams understands the morality system.

What is most fundamental to the morality system is, Williams argues, its “special notion” of moral obligation. There is, he says, a range of ethical outlooks that adopt this idea of moral obligation (Williams 1985/2011: 93-4). Moral obligations are, for these ethical outlooks, grounded in reasons that are available to all rational agents. These reasons serve as “practical necessities” for the agents concerned and give over-riding weight to “moral” over “non-moral” considerations (Williams 1985/2011: 203-05, 209). These obligations apply to us all as members of the moral community or “the universal constituency” as Williams describes it (Williams 1985/2011: 16). “Morality’s” (special) notion of obligation is closely bound up with related concepts of voluntariness and blame. Agents who voluntarily violate the demands of morality are subject to blame and retribution (Williams 1985/2011: 200; Williams 1993/1995: 72-4). When we violate our moral obligations we *deserve* to suffer as a matter of justice (Williams 1985/2011: 214; Williams 1993/1995: 72; Williams 1994: 8). Although Kant is the principal representative of this outlook, it includes others, such as utilitarianism (Williams 1985/2011: 194-8). More generally, however, this outlook “is not an invention of philosophers. It is the outlook, or, incoherently, part of the outlook of almost all of us” (Williams 1985/2011: 195).

What motivates this conceptual apparatus is an aspiration to ensure “the purity of morality”, which requires, most importantly, that morality must be immune to the influence of luck (Williams 1985/2011: 216-8). This is essential if morality is to satisfy its ideal of “ultimate justice” (Williams 1985/2011: 218). In order to support this idealized picture of the human moral community we need much more than the mundane elements of psychological explanation, such as belief, desire, intention, deliberation, and so on (Williams 1993: 31-4, 46-7, 40, 46, 55-6, 66-7). Morality, Williams suggests, must “deepen” and “refine” the idea of the voluntary in order to enable us to “transcend luck” (Williams 1985/2011: 215-7; and also Williams 1976/1981).[[19]](#footnote-19) To make the idea of the voluntary “profound” what is required is “total control” or “ultimate freedom” of some kind (Williams 1985/2011: 63-5; Williams 1993: 158; Williams 1994: 7-9), which is impossible to satisfy or even make sense of (Williams 1993/1995: 44). For “morality” it is, nevertheless, essential that we are able to deepen the idea of the voluntary since, failing this, there is no hope of “transcending luck” or securing “ultimate justice”.[[20]](#footnote-20)

These “illusions” about moral responsibility, and the associated effort to “deepen” the voluntary, Williams argues, that generate the free will problem:

... there is a problem of free will only for those who think that the notion of the voluntary can be metaphysically deepened. In truth, though it may be extended or contracted in various ways, it can hardly be deepened at all. What threatens it is the attempt to make it profound, and the effect of trying to deepen it is to put it beyond all recognition. [[21]](#footnote-21)

Williams is careful to stress that efforts to provide depth to the concept of responsibility are not the sole property of libertarian theories. On the contrary, most compatibilists are no less committed than their libertarian counterparts to “keeping the morality system in adequate business” (Williams, 1985/1995: 19). While Williams rejects any general scepticism about moral responsibility, the compatibilism that he defends is one that requires us to “recast our ethical conceptions” relating to responsibility and blame. Williams’ compatibilism does not aim to “leave everything where it was” (Williams 1995: 578).[[22]](#footnote-22)

This brings us to the question of whether or not Wallace’s reflective self-control model can satisfy the demands of “morality”? There is, I think, reason to doubt this. What we still face may be described as the limits of agency as it concerns free will and moral responsibility. Agents who satisfy compatibilist conditions of “normative competence” of the kind described by Wallace may possess powers of reflective self-control but the specific dispositions that they acquire and come to possess will be very largely shaped by conditions they have no control over and cannot influence.[[23]](#footnote-23) Moreover, once acquired, although they may well exercise these powers in various ways, they have no ability to control how they are exercised in the particular circumstances they may encounter or be placed in. (It is worries of this nature that motivate the “bipartite accounts” that Wallace considers and rejects.) Beyond this, what specific circumstances the agent does or does not find herself in will also depend very largely on factors that she has little or no control over. All these considerations plainly apply to JoJo and those like him, even though JoJo (per hypothesis) may otherwise be entirely “normatively competent” as judged by the standard of the reflective self-control. The concern lying behind this is that, in the final analysis, the particular moral trajectory that an agent’s life will take depends to a considerable extent on factors that they have little or no control over - *despite* possessing and exercising powers of reflective self-control. Considered from the perspective of “the morality system”, this conclusion is liable to lead to the need for “autonomy” of some metaphysical kind. From there it is a short step to plain scepticism about moral responsibility.

***IV. The Reason View and “Morality****”*

We have noted that although Wallace’s reflective self-control account shares much with Wolf’s Reason view as it concerns “normative competence”, it diverges from it in respect of symmetry and the principle of alternative possibilities (as this relates to blame). We have also noted that Wallace’s reflective self-control model explicitly claims to satisfy the requirements of “the morality system” and its understanding of responsibility and blame. Given this, we may now ask the following two questions:

(1) Does the Reason View (also) aim to satisfy “morality” and its (“peculiar”) conception of blame?

(2) If so, does the Reason View succeed in this ambition where the reflective self-control account fails?

It should be noted, from the beginning, that not only is Wolf a sympathetic and illuminating commentator on Williams’ ethics, several of her studies suggest that she sides with Williams on significant matters relating to his criticisms of “morality”.[[24]](#footnote-24) Nevertheless, for the purpose of the present discussion, our focus should remain on the Reason View and its commitments. In respect of praise (and deep responsibility), it may be argued that the Reason View diverges from “morality’s” requirement that desert presupposes alternative possibilities.[[25]](#footnote-25) On the other hand, with regard to deserved blame, Wolf’s appeal to the principle of alternative possibilities as an essential condition of responsibility for wrongdoing, has the appearance of being motivated by “morality’s” aspiration to “transcend luck” and ensure conditions of (absolute) fairness or “ultimate justice”.[[26]](#footnote-26)

Although the Reason View rejects any effort to secure *depth* by means of some (metaphysical) account of “ultimate control”, it is still concerned to provide an alternative account of depth (FWR, 5, 10-1,42, 44-5, 50-5,63-6, 71-3, 75-6, 77, 131).[[27]](#footnote-27) Without depth, as provided by Reason and our ability to do the right thing for the right reason, no agent would *deserve* praise or blame. The existence of agents, such and JoJo and those like him, is evidence that some of those who do wrong or act badly have no *opportunity* to do the right thing for the right reasons. Due to their deprived and damaged background histories, these agents cannot appreciate the True and the Good. In the absence of any relevant opportunity to act otherwise, the fact that these agents simply possess and exercise powers of rational self-control will not suffice for deserved *blame*. The Reason View aims to fill this gap by means of the additional “negative” requirement that nothing “blocks the exercise of one’s relevant capacities and skills” (FWR, 101-02). It is not evident, however, that this “negative” condition is itself sufficient to “transcend luck” or secure any form of “absolute fairness” or “ultimate justice”.

Even if an agent is not “interfered” with or “prevented” from exercising their rational powers it is not obvious that this ensures that the agent is not subject to fate and luck of some relevant kind. An agent such as Rose, who satisfies Reason View conditions, still *acquires* the specific dispositions and rational capacities that she has in ways that she has little or no control over (e.g. generally they are acquired early in childhood). Moreover, although Rose may *exercise* those capacities without them being “blocked” or subject to causal deviant factors of some kind (i.e. as we find in cases involving coercion, hypnotisms, manipulation, and so on: FWR, 105,108), she still lacks any effective *exercise control*, since how her capacities are actually exercised is not itself something that she controls. (It is this form of further control that the Autonomy View aims to deliver.) When Rose fails to do what Reason requires (e.g. grade her papers) this failure may still be explained in terms of factors that she has no control or influence over. The general worry here is that the line that Wolf aims to draw between cases like JoJo and Rose is not entirely convincing. It may, for example, be argued that the background “histories” of agents who do what is wrong or act against Reason, may *all* have some relevant (if often unknown) explanation for their failures (e.g. childhood deprivation, poor moral education, troubled family relationships, faulty cultural prejudices, lack of sleep or stress, etc.). In that case, excusing considerations of this kind may be generalized to all cases of wrongdoing (e.g. Rose) and the sceptical threat returns. [[28]](#footnote-28)

The fundamental problem with the Reason View, interpreted this way, is that it aims to satisfy “morality’s” demand for “purity” and “absolute justice” but it lacks the (metaphysical) resources to do this - because no coherent account can deliver this. This returns us to the threat of scepticism about (deep) responsibility. If the Reason View aims to satisfy “morality” in this respect then, for reasons similar to those that applied to Wallace’s reflective self-control model, it must be judged a failure. In response to this, it may be suggested that, in contrast with Wallace’s aims, the Reason View does not aim to satisfy “morality” in these respects. However, if that is the case, then it becomes unclear why the Reason View adopts the apparatus of asymmetry and alternate possibilities. Not only are these features problematic to make sense of and apply, they fail to ensure that the responsible agent remains immune to fate and luck in the manner that “morality” requires.

The above analysis of the relationship between the Reason View and “the morality system” suggests that the Reason View - and, perhaps, all compatibilist-oriented rational self-control accounts like it - face a basic dilemma. On one hand, we may take the Reason View to aim to satisfy the core requirements of “morality” and its particular understanding of what deep responsibility demands. If that is the case, then, as we have argued, the Reason View and accounts like it, fail. On the other hand, if the Reason View does not aim to satisfy “morality” in these terms, then not only is the apparatus of asymmetry and alternate possibilities unnecessary (and needlessly problematic), the position defended cannot be presented as addressing “the free will problem” as it is generally understood. Doing that requires saving *deep* responsibilit*y* *as “morality” understands it* from sceptical threat. Abandoning the free will problem, understood in the terms “morality” suggests, would, in any case, run contrary to how Wolf presents the Reason View. The core aim of the Reason View is to provide credible and coherent depth of the kind that “morality” requires (FWR, 42-5, 66, 90-3; SMR, 281-2). Since the Reason View is not in the business of “revising”, much less “recasting”, the ethical concepts involved, any project of a more radical kind - dispensing with “morality” - is not something that Wolf’s approach would encourage.

The contrast that we have drawn between the Reason View and Williams’ critique of “the morality system” suggests that there is considerable distance between them. Although the Reason View rejects the efforts of the Autonomy View to secure depth in terms of “metaphysical” powers of ultimate control, it still seeks to satisfy the requirements of “depth” that the morality system is committed to (i.e. to “transcend luck” and secure “ultimate justice” or “absolute fairness” of some kind). Williams maintains that this cannot be done. This is, of course, what the sceptic also maintains with respect to any conception of responsibility interpreted along these lines (i.e. in terms of “accountability”, “basic desert”, “true responsibility”, etc.). Where Williams diverges from the sceptic is that he aims to *vindicate* responsibility by suggesting that we need to “recast” concepts in more “realistic” and “truthful” terms. The position and methods that he suggests may be described as “responsibility realism”.[[29]](#footnote-29)

***V. Hard Cases and Responsibility Realism***

Proper examination of responsibility realism and what it does or might involve would take us well wide of our immediate concerns in this context, which are focused on the Reason View. There is, however, one important feature of the Reason view that we should still consider as it relates to responsibility realism. This concerns the relevance of Wolf’s “hard cases” (i.e. JoJo et al) for our understanding of the boundaries of the moral community, which serve to identify those agents who we take to be responsible. One feature of the Reason View that responsibility realism can accept and endorse is that we must identify those who are responsible agents with reference to rational abilities of some kind. Apart from anything else, it is the absence of these general abilities and capacities that serve to *exclude* animals, young children, the insane, and others like them, from the moral community.[[30]](#footnote-30) To this extent the responsibility realist retains important insights as provided by the Reason View and its correlates. The extent of this agreement, however, should not be exaggerated.

Just as the responsibility realist is sceptical about any effort to appeal to powers of Reason to escape or “transcend luck” within our ethical lives, so too the responsibility realist is sceptical about the (mistaken) idealism involved in supposing that the moral community is constituted in such a way that all those within it can come to agree about what moral reasons there are - and that any failure to do this is evidence of a failure of rationality.[[31]](#footnote-31) The difficulty that we face in respect of this, however, is that we routinely encounter agents who are “normatively competent” in all relevant respects - and certainly “rational” in terms of their ability to operate and function in human society - but who we are, nevertheless, still unable reach or influence (i.e. “correct”) through *reasoning* of any kind.[[32]](#footnote-32) What these agents lack, according to the responsibility realist, is any relevant subjective disposition that can be engaged or provide sufficient motivation to give salience and significance to ethical considerations of some kind.[[33]](#footnote-33) Whereas the adherents of morality system (e.g. the Reason View) would place these individuals (e.g. as described in Wolf’s “hard cases”) *outside* the boundaries of the moral community, the responsibility realist would draw the boundary differently.

Wolf’s example of JoJo helps to make clear what the difference between the Reason View and responsibility realism comes to in respect of this. How, we may ask, would the responsibility realist understand JoJo and those like him (e.g. Nazis, et al)? These are individuals who have done wrong and acted badly but who, Wolf argues, we do not regard as responsible. Due to their background histories (i.e. deprivation, faulty moral education and upbringing, etc.) they have acquired “value systems” that they are unable to “correct” or shed (FWR, 33, 112-3; SMR, 287-9). While JoJo and those like him may enjoy “capacities, talents, and skills” that differentiate them from animals, young children and the insane, they are still, according to the Reason View, insufficiently competent in these respects to be responsible or an appropriate targets of blame (i.e. as judged in terms of the requirements of “deep responsibility”).

From the realist perspective it *matters* that JoJo and those like him are, unlike animals, young children, or the insane, able to participate and engage in human *ethical* life. More specifically, however ethically damaged these individuals may be in respect of their “valuational systems”, they are, nevertheless, still able to understand moral norms and expectations, as well as the various moral reactions and responses that they may give rise to (i.e. reactive attitudes of various kinds). Clearly they do not share our values. Nor can they “correct” their values or conduct in light of various relevant (ethical) considerations that we may present to them. In these respects, therefore, they cannot be reached through rational means.[[34]](#footnote-34) It does not follow from this, however, that they are unable to comprehend our norms and values and why we react and respond to their (wrongful) conduct as we do. While it may be pointless to expect or hope that these individuals can be “corrected” or “reformed” by means of blame and negative reactive attitudes of various kinds, it would be untruthful to deny that these individuals are not only as human agents but that they are *ethical* agents.[[35]](#footnote-35) Whatever the limits of reason may be with respect to our dealings with “hard cases” of this kind, viewing them as responsible agents is a matter of recognizing that they are *not* simply objects to be “treated” and “trained”.[[36]](#footnote-36) On the contrary, however ethically damaged they may be, they are individuals who still possess (ethical) capacities and qualities that require us to regard and respond to them in fundamentally different terms than this. Clearly there is nothing “superficial” about regarding these individuals in this way.

In sum, the contrast between the (morality-oriented) Reason View and responsibility realism in relation to hard cases of this kind can be understood in terms of their contrasting understanding of the way in which (hard case) individuals such as JoJo are taken to be *victims* of their unfortunate formative backgrounds. For the Reason View their ethical misfortune consists in being rendered incapable of responsible agency (i.e. because they are incapable of correcting or shedding their corrupt and distorted valuational systems). As such, they must be viewed as subjects of “treatment” or “therapy”. For the responsibility realist, the misfortune that these individuals suffer is that they are blameworthy agents whose conduct is, nevertheless, plainly subject to fate and luck in respect of the particular (catastrophic) ethical trajectory that their lives have taken. For the responsibility realist, it is a troubling truth about the human ethical predicament that this is a form of moral misfortune that is all too prevalent in human society and history.[[37]](#footnote-37)

**Bibliography**

Dennett, Daniel. (1984). *Elbow Room: The Varieties of Free Will Worth Wanting*. Oxford: Oxford University Press.

Ekstrom, Laura. 2000. *Free Will: A Philosophical Study*. Boulder, CO: Westview Press.

Fischer, John M. & Mark Ravizza. 1998. *Responsibility and Control*. Cambridge: Cambridge University Press.

Frankfurt, Harry. 1971. “Freedom of the Will and the Concept of a Person”. *Journal of Philosophy* 68: 5-20.

McKenna, Michael, and P. Russell, eds. 2008. *Free Will and Reactive Attitudes: Perspectives on Strawson’s “Freedom and Resentment”*. London & New York: Routledge.

Nagel, Thomas. (1976/2013). “Moral Luck”. *Proceedings of the Aristotelian Society*, Supp. Vol 1 (1976), 137-51. Reprinted in P. Russell & O. Deery, eds. *The Philosophy of Free Will*:31–42.

Nelkin, Dana. 2011. *Making Sense of Freedom & Responsibility*. Oxford: Oxford University Press.

- 2013. “Prècis of *Making Sense of Freedom and Responsibility*”. *Philosophy and Phenomenological Research* 86: 443-458.

Pereboom, Derk. 2014. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.

Queloz, Matthieu. 2022. “A Shelter from Luck: The Morality system Reconsidered”. In A. Szigeti and M. Talbert, eds. *Morality and Agency: Themes from Bernard Williams*. New York: Oxford University Press; 182-209.

Reid, Thomas, 1788/1969. *Active Powers of the Human Mind*. Cambridge, MA: MIT Press.

Russell, Paul. 2017. “Free Will Pessimism”. In D. Shoemaker, ed. *Oxford Studies in Agency and Responsibility*, Vol. 4. Oxford University Press. Pp. 93-120. Reprinted in Russell, *The Limits of Free Will*; 243-75.

- 2018. “Bernard Williams: Ethics from a Human Point of View”. *Times Literary Supplement* (18 December)

- 2022. “Free Will and the Tragic Predicament: Making Sense of Williams”. In A. Szigeti and M. Talbert, eds. *Morality and Agency: Themes from Bernard Williams*. New York: Oxford University Press; 163-81.

- [Forthcoming.] “Recasting Responsibility: Hume and Williams”. In Marcel van Ackeren & Matthieu Florian Queloz, eds.), *Bernard Williams on Philosophy and History*. Oxford: Oxford University Press.

Russell, Paul, and Oisin Deery, eds. 2013. *The Philosophy of Free Will: Essential Readings from the Contemporary Debates*. New York & Oxford: Oxford University Press.

Strawson, Galen. (1994/2013). “The Impossibility of [Ultimate] Moral Responsibility”. *Philosophical Studies*, 75 (1994), 5-24. Reprinted in P. Russell & O. Deery, eds. *The Philosophy of Free Will*; 363-78.

Taylor, Charles. 1976. “Responsibility for Self”. In *The Identities of Persons*. A.O. Rorty, ed. Berkeley: University of California Press. Reprinted in G. Watson, ed. *Free Will*. Oxford: Oxford University Press, 1982: 111-126

Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.

Watson, Gary. 1975. “Free Agency”. *Journal of Philosophy* 72: 205-220. Reprinted in G. Watson, ed. *Free Will*. Oxford: Oxford University Press, 1982: 96-110.

- 2004. “Two Faces of Responsibility”. In G. Watson, *Agency and Answerability: Selected Essays*. Oxford: Oxford University Press, 2004: 260-288.

- 2013. “Asymmetry and Rational Ability”. *Philosophy and Phenomenological Research* 86: 467-475.

Williams, Bernard. 1976/1981. “Moral Luck”. *Proceedings of the Aristotelian Societ*y, Supp. Vol. 1 (1976), 115-35. Reprinted in B. Williams, *Moral Luck: Philosophical Papers 1973-1980*. Cambridge: Cambridge University Press: 20-39.

- 1980/1981. “Internal and external reasons”. In R. Harrison, ed. *Rational Action*. Cambridge: Cambridge University Press, 1980. Reprinted in B. Williams, *Moral Luck: Philosophical Papers 1973-1980*. Cambridge: Cambridge University Press: 101-113.

- 1985/2011. *Ethics and the Limits of Philosophy*. With a commentary by A.W. Moore and a foreword by J. Lear. London & New York: Routledge.

- 1985/1995. “How free does the will need to be?”. Lindley Lecture, University of Kansas (1985). Reprinted in B. Williams, *Making sense of humanity* Cambridge: Cambridge University Press: 3-21.

- 1993/1995. “Nietzsche’s minimalist moral psychology”. *European Journal of Philosophy* 1.1 (1993): 4-14. Reprinted in Williams, *Making sense of humanity* Cambridge: Cambridge University Press: 65-76.

- 1993. *Shame and Necessity*. Berkeley, CA: University of California Press.

- 1994. “Interview with Bernard Williams”. *Cogito*, 8.1: 3-19.

- 1995. “Ethics”, *Philosophy: A Guide Through the Subject*. Ed. A.C. Grayling. Oxford: Oxford University Press.: 545-582.

Wolf, Susan. 1980.“Asymmetrical Freedom”. *Journal of Philosophy* 77: 157-66.

- 1981. “The Importance of Free Will”. *Mind* 90: 386-405. Reprinted in M. McKennna & P. Russell, eds. *Free Will and Reactive Attitudes*: 69-83.

- 1982. “Moral Saints”. *Journal of Philosophy* 79: 419-439.

- 1987. “Sanity and the Metaphysics of Responsibility”. *In Responsibility, character, and the emotions: New essays in moral psychology*. F.D. Shoeman, ed. Cambridge: Cambridge University Press: 46-62. Reprinted in P. Russell & O. Deery, eds. *The Philosophy of Free Will*: 281-295.

- 1990. *Freedom Within Reason*. Oxford: Oxford University Press.

- 2012. “’One Thought Too Many’: Love, Morality, and the Ordering of Commitments”. Reprinted in S. Wolf, *The Varieties of Values: Essays on Morality, Meaning & Love*. Oxford: Oxford University Press: 143-162.

ABSTRACT: This paper examines Susan Wolf's account of "the Reason View" of moral responsibility as articulated and defended in *Freedom Within Reason* (OUP 1990). The discussion turns on two questions about the Reason View: (1) Does the Reason View aim to satisfy what Bernard Williams describes as “morality” and its (“peculiar”) conception of responsibility and blame? (2) If it does, how successful is the Reason View judged in these terms? It is argued that if the Reason View aims to satisfy “morality” in respect of its understanding of deserved blame -- as it seems to -- then it fails for reasons similar to those that apply to R. J. Wallace’s “reflective self-control” model. On the other hand, if Wolf’s Reason View does not aim to satisfy “morality” in this respect then it might well rest satisfied with the more limited conditions of Wallace’s Rational Self-Control model, which in contrast with the Reason View, does not appeal to the (problematic) apparatus of asymmetry and the principle of alternate possibilities.

1. I am grateful to the audience attending the Lauener Prize awards ceremony, celebrating Susan Wolf’s work (in Bern in August 2022) for their interesting comments and discussion. I would also like to thank the audience attending the Erlangen Workshop on “Moral Responsibility in a Strawsonian Key” (June 2024) for their similarly helpful comments and discussion on a later draft. [↑](#footnote-ref-1)
2. Wolf, *Freedom Within Reason*, 45. [↑](#footnote-ref-2)
3. The paper that is especially relevant to the position that Wolf carves out in FR is “Sanity and the Metaphysics of Responsibility” (1987), hereafter abbreviated as SMR. See also, however, “Asymmetrical Freedom” (1980), and “­The Importance of Free Will” (1981). [↑](#footnote-ref-3)
4. Among the most important and influential of these are Dennett, 1984; Wallace, 1994; Fischer & Ravizza, 1998. [↑](#footnote-ref-4)
5. More recently, Wolf’s asymmetrical view has been modified and defended by Dana Nelkin, in what she refers to as the “rational abilities” view (Nelkin, 2011; see also Nelkin, 2013). Given that there are a number of affinities between Wolf’s Reason View and Nelkin’s “rational abilities view”, some of the concerns raised in this paper concerning the Reason View could also be applied to Nelkin’s rational abilities view. [↑](#footnote-ref-5)
6. “... the concept of responsibility is a mysterious one which tends, on examination, to become increasingly opaque and to threaten variously to be incoherent or impossible or universally inapplicable.” (FWR, 3) [↑](#footnote-ref-6)
7. The concept of “deep responsibility” that Wolf is concerned with has been given different labels in the context of the contemporary the free debate. Some, for example, refer to it as “accountability” (Watson, 2004: 262-7), others as “basic desert” (Pereboom, 2014: 2) or “true responsibility” (Strawson, 1994/2013: 363). Whatever label we choose, however, it is this core conception that serves as the focus of the free will debate. [↑](#footnote-ref-7)
8. Included among the theorists who defend the Real Self View are Frankfurt, 1971; Watson, 1975; and Taylor, 1976. [↑](#footnote-ref-8)
9. In FWR there is a (terminological) variation in Wolf’s way of describing JoJo type cases. In SMR JoJo is said to have a “deep self” – but lacks a *sane* deep self (i.e. because he and others like him are “governed by mistaken conceptions of value that [they] cannot help but have” (SMR, 290). In FWR Wolf draws a contrast between “the real self” and “the deep self”. In problematic cases (i.e. JoJo, Tony, et al), while the agent may have a “real self” they lack a *deep* self - where this is the kind of real self that would license blaming them for their actions. (FWR, 43-4). What we need, on this analysis, is an explanation of why these “real selves” are not sufficiently *deep* selves. [↑](#footnote-ref-9)
10. Wolf’s claim that these individuals are responsible only in a “superficial” sense is challenged in Watson, 2004: 268-273. Watson argues that the kind of responsibility involved in cases of this kind is “aretaic” or made from an “aretaic perspective”. He describes this as an “attributability” conception and argues that it is “a core notion of responsibility”, even though he agrees that it needs to be distinguished from “accountability” conceptions of the (deep) kind that Wolf is concerned to vindicate. [↑](#footnote-ref-10)
11. In support of this claim Wolf gives the example of a swimmer who is prompted by Reason to save a drowning child. It is not evident, she argues, that we think that agent who chooses to save the child would be more free or more responsible in any *worthwhile* sense if they could not have chosen otherwise in this situation (FWR, 57-9). Cases like this show, it is claimed, “that responsibility does not depend on autonomy” (FWR, 61) [↑](#footnote-ref-11)
12. In a discussion of Nelkin’s similar position, Gary Watson comments on asymmetry in rather sharper terms: “To call Asymmetry ‘unorthodox’ is an understatement. Wolf and Nelkin are, as far as I know, its sole adherents...” Watson, 2013: 468. [↑](#footnote-ref-12)
13. We might contrast this claim with Thomas Reid’s contrary view, which is representative of the Autonomy View. Reid comments on the Roman hero Cato: “What was, by an ancient author, said of Cato, might indeed be said of him. He was good because he could not be otherwise. But this saying, if understood literally and strictly, is not the praise of Cato, but of his constitution, which was no more the work of Cato, than his existence” (Reid, 1788/1969: 261). As Reid sees it, understood this way, Cato no more deserves praise than JoJo deserves blame. [↑](#footnote-ref-13)
14. Fischer and Ravizza argue against Wolf’s asymmetrical view on the ground that it is a mistake to suppose that *all* cases where an agent cannot do the right thing can be assimilated to cases where the agent’s ability to recognize and respond to reason is in some way damaged or impaired (Fischer and Ravizza, 1998: 55-61). What exempts an agent from responsibility in cases involving addiction, kleptomania, and so on, is not that the agent could not have done otherwise but that their relevant rational powers and abilities are undermined or impaired. (For a related discussion see Wallace, 1994: 166-82.) [↑](#footnote-ref-14)
15. Wallace argues that “Wolf apparently assumes that a condition should count as *preventing* the exercise of our general powers only if it is sufficiently like such paradigm cases as those involving physical constraint, paralysis, or automatic muscular contraction.... But incompatibilists will ask why the physical impossibility of exercising our general powers is not sufficiently like these paradigm cases to be counted an opportunity-defeating condition as well” (Wallace, 1994: 209-10). [↑](#footnote-ref-15)
16. Wolf concedes, however, that in practice it is often “uncertain” how “difficult” it was for the agent to have done otherwise. To this extent “the Reason View implies that we are not always in a position to know whether an agent is responsible and blameworthy” (FWR, 86-7, 121-2). Critic may argue, more strongly, that once we allow for concerns about psychological determinism along these lines, we are *never* in a position to know whether an agent is responsible and blameworthy (i.e. there may be all kinds of facts about Rose’s particular psychology that she/we are unaware of that may fully explain her wrong action). [↑](#footnote-ref-16)
17. Wallace, 1994: 202, 206-07, 215. Nelkin presents and defends her asymmetrical, “rational abilities” view in terms of considerations of fairness and the principle of alternative possibilities. See, especially, Nelkin, 2011: Chp. 2. [↑](#footnote-ref-17)
18. Wallace acknowledges that there may be (background) circumstances in which it may be “extremely difficult” for a person to grasp and comply with moral demands of a given kind and suggests that considerations of this kind may serve to mitigate blame or exempt (Wallace, 1994: 214-6, 232). [↑](#footnote-ref-18)
19. On this reading, the demand for depth is itself an expression of “morality’s” aspiration to “transcend luck” and secure “ultimate justice”. The classic presentation of this demand can be found in Nagel’s presentation of the problem of moral luck in Nagel 1976/2013. [↑](#footnote-ref-19)
20. The crucial psychological element required for this - which “morality” introduces and places emphasis on - is “the will” (Williams 1993: 5, 28-37, 46). [↑](#footnote-ref-20)
21. Williams 1993: 68. [↑](#footnote-ref-21)
22. For a general overview of Williams’ views on free will and moral responsibility see Russell (2022). See also, on Williams’ criticism of “the morality system”, Russell (2018) and Queloz (2022). [↑](#footnote-ref-22)
23. With respect to this, the familiar facts of moral development and education make clear that the scope for amending (or improving) moral character shrinks rapidly early in life - well before agents reach an age of maturity or acquire the critical means necessary to fundamentally reform themselves by means of reflection. To the extent that this *does* remain possible at a later stage, it will largely depend on (other) prior conditions that the agent has little control over. [↑](#footnote-ref-23)
24. Perhaps the most important of these concern Williams’ criticism of the assumption that *moral* concerns (always) have priority over *non-moral* concerns. On this see, in particular, Wolf, 1982; and also Wolf, 2012. [↑](#footnote-ref-24)
25. Consider, again, the remarks of Reid on Cato’s virtue, as cited further above (note 13). See. Also Ekstrom 2000: esp. 164-69 for a forceful statement of the Autonomy View perspective on this matter. [↑](#footnote-ref-25)
26. The first paper that Wolf cites in *Freedom Within Reason* is Nagel’s classic discussion of “Moral Luck” (FWR, 149n1; citing Nagel 1976/2013). Nagel’s paper is cited in the context of a discussion concerning the threat of scepticism about responsibility (FWR,10). In what follows, however, Wolf says little or nothing about the problem of moral luck as it specifically concerns the sceptical threat. While it seems reasonable to suppose that the problem of scepticism and of moral luck overlap, they need to be differentiated, since not everyone will agree that the existence of moral luck implies scepticism. Beyond that, the various particular modes of moral luck also need to be distinguished, since there may be some debate about *which* modes of moral luck pose a sceptical threat. [↑](#footnote-ref-26)
27. It is, of course, significant that the Reason View is in agreement with the Autonomy View’s assessment that the Real Self View is “superficial” and fails to account for *deep* (deserved) praise and blame. According to the Reason View, however, what provides the sort of *depth* that we are looking for is not to be found in the metaphysics of “ultimate control” but in the normative powers of Reason and the “special power or control” that it provides (FWR, 6, 76-7, 92-3). [↑](#footnote-ref-27)
28. Wolf’s recognition that the ability to act otherwise is “a matter of degree” and not an “all-or-nothing matter” adds fuel to this sceptical concern, since blame, on close examination, may *always* dissipate or bleed away (FWR, 86-7). [↑](#footnote-ref-28)
29. Although this is not Williams’ own terminology it is, nevertheless, faithful to his concerns and his other remarks. The vindication of responsibility realism, so understood, is the primary task of Williams’ *Shame and Necessity*. This project may be understood, in more general terms, as aiming to provide a more “truthful” account of moral responsibility, one based on a more realistic moral psychology than that which “morality” relies on (i.e. minus the illusions relating to *pure* and *ultimate* agency, etc.). On this see, in particular, Williams 1993/1995. [↑](#footnote-ref-29)
30. We exempt these individuals from moral evaluation and the responses associated with them (e.g. reactive attitudes of various kinds) because they are incapable of interpreting and internalizing moral norms or formulating moral values and acting according to them. These are activities and modes of agency that require rationality - as filtered through particular languages and cultures that evolve over historical time. [↑](#footnote-ref-30)
31. Wolf is careful to emphasize that while the Reason View is committed to some notion of “the True and the Good”, it leaves scope for “Normative Pluralism” (FWR, 118, 135-42). What is crucial to this view is that “the metaethical assumptions of the Reason View are quite weak” (FWR, 118, 123). In particular, while it is committed to “the existence of nonarbitrary standards of correctness for value judgments” and has some sort of “commitment to the objectivity of value”, this “need not determine a unique universally applicable, complete and optimal system of values and value judgments” (FWR, 124; and cp. 134-5). However, even allowing for “Normative Pluralism”, the responsible agent must still “be capable of forming better values rather than worse ones, good value judgments rather than bad ones, just insofar as there *are* better and worse choices and judgments to be made” (FWR, 125 - Wolf’s emphasis). Any agent who is incapable of this (e.g. JoJo) cannot be a responsible agent. [↑](#footnote-ref-31)
32. As Williams points out, someone who violates obligations or the rights of others may “not necessarily [be] behaving irrationally or unreasonably, but badly. We cannot take for granted that he had a reason to behave well, as opposed to our having various reasons for wishing that he would behave well...” (Williams 1985/2011: 213-4). I take this to show that, however bad or evil JoJo may be, he not necessarily *irrational* - since he may have no reason to do what is right. [↑](#footnote-ref-32)
33. The fundamental worry here concerns the limits of practical reason itself and the way in which is embedded in “an agent’s subjective motivational set” (Williams 1980/1981: 102). The *limits* of reason - and of (philosophical) “ethical theory” in particular - as it concerns giving direction and shape to ethical life, is a major theme in Williams’ *Ethics and the Limits of Philosophy*. [↑](#footnote-ref-33)
34. These individuals may, of course, like young children, animals and the insane, be “reached” by means of the threat of “sanctions” of some kind, and this may well alter their actual conduct. [↑](#footnote-ref-34)
35. What is especially important is that they have sufficient rational and psychological abilities to *understand* the valuational systems of others, even though they may reject or repudiate them. This is something that animals, young children, and the insane cannot do, which is why we place them outside the boundaries of the moral community. From this perspective it is a mistake - as encouraged by “morality” - to limit the boundaries of the moral community to individuals who either share our values, or can in principle, come to do so (i.e. who we can always be “reached through reason”). [↑](#footnote-ref-35)
36. Suffice it to point out that, for the responsibility realist, it is still true that many, if not most, individuals who act wrongly and deserve blame are not beyond reasoning and can be influenced and “corrected” through these channels. Prevalent as “hard cases” like JoJo may be, not *all* cases of blame or wrongful action should be assimilated to this. Nevertheless, the responsibility realist still considers these (hard case) individuals to be responsible agents and members of the moral community (whereas the Reason View does not). [↑](#footnote-ref-36)
37. For further elaboration on these and related themes and concerns see Russell (2017) and Russell (forthcoming). [↑](#footnote-ref-37)