

# Narcissism, Entitlement, Responsibility

Forthcoming in the *Australasian Journal of Philosophy*.

Hannah Altehenger

University of Konstanz, Fachbereich Philosophie, Universitätsstraße 10, 78464 Konstanz  
Germany

**Abstract:** Recent years have seen a surge of interest in the topic of moral responsibility for 'non-ideal' agents. And yet, one important type of 'non-ideal' agent, the narcissistic agent, has not received much attention. In this paper, I seek to fill this gap. My focus is on psychological entitlement, a feature that has been largely overlooked. I argue that this feature impairs narcissistic agents' moral competence. This is because it both causes them to form distorted moral assessments in a wide range of situations and impairs their ability to use feedback from others to correct these distortions. I conclude that narcissistic agents have mitigated responsibility owing to their impaired moral competence. As I furthermore show, this does not entail that we simply need to accept the damage they do. Rather, we may take steps to protect ourselves against the destructive effects of narcissistic entitlement, both on a personal and on a societal level.

**Key words:** narcissism; responsibility; accountability; 'non-ideal' agency; moral competence requirement; entitlement

## 1. Introduction

A party guest dominates the conversation and will not stop bragging about their yacht vacation.

A customer at the agency repeatedly jumps the queue. A fellow traveller makes long phone calls while sitting in the quiet zone. A friend calls you at all hours for lengthy talks about their problems but when you need someone to talk to is nowhere to be found. A parent leaves the concert hall during their child's solo performance because they cannot bear seeing their child succeed. A child uses their parent's funeral as a stage to showcase their own poetry. A co-worker takes all the credit for your mutual work, thereby securing themselves a promotion. A CEO keeps rejecting excellent ideas because they are not their own and thereby hurts the company. A politician embezzles huge amounts of public welfare money in order to finance their luxury lifestyle.

These examples illustrate that narcissism can manifest in many ways and that its manifestations often annoy, enrage or hurt us. It seems natural to react to this kind of behaviour

with resentment or indignation. ‘How dare you behave like that!’ we might want to shout at more than one of the agents just described. We might even feel that imposing punitive measures is appropriate in some of the examples.

In short, narcissistic agents often display behaviour for which we want to hold them accountable and for which we believe they should be held accountable. This paper asks whether our pre-theoretical stance toward narcissistic agents’ responsibility is justified.

The paper should fill an important gap. Even though recent decades have seen a surge of interest in the topic of moral responsibility for ‘non-ideal’ agents (see, e.g., Kennett 2002; Jaworska 2007; Watson 2011; Shoemaker 2017), the narcissistic agent has not yet received much attention.<sup>1</sup>

The key thesis for which I shall argue is that our pre-theoretical stance toward narcissistic agents’ responsibility is *not* justified: Close inspection reveals that narcissism is a condition that mitigates moral responsibility, and, in some cases, even significantly so. More specifically, I shall argue that narcissistic agents fail to (fully) fulfill the *moral competence requirement* for responsibility (Talbert 2024: § 3.2). This is because of a certain deep-seated feature of the narcissistic personality that has been called *psychological entitlement* (Campbell, Bonacci, *et. al.* 2004; see also Miller, Hoffman, *et. al.* 2011; Krizan and Herlache 2017). This feature, I maintain, causes narcissistic agents to form distorted moral assessments in a wide range of situations, while simultaneously undermining their ability to use feedback from others to correct these distortions.

My view entails that reacting to narcissistic agents’ bad behaviour with a (full) display of robust blaming responses, such as expressed resentment, expressed indignation, or various forms of adverse treatment is inappropriate. That said, we need not simply accept the damage narcissistic agents cause. As I will furthermore show, we may take various steps to protect ourselves against the destructive effects of their entitlement.

---

<sup>1</sup> There are a few recent exceptions, in particular Fatic (2023a; 2023b) and Shoemaker (2024). I will come back to their positions below (see section 3).

My paper is organized as follows. In Section 2, I introduce the moral competence requirement. In Section 3, I provide some empirical background on narcissism, with a special focus on psychological entitlement. In Section 4, I argue that psychological entitlement impairs narcissistic agents' moral competence, and, thereby, mitigates their responsibility. In Section 5, I defend my view against objections.

## 2. The Cognitive Moral Competence Requirement

I shall rely on a standard conception of moral responsibility as *accountability*.<sup>2</sup> On that conception, an agent is morally responsible (blameworthy) for their morally problematic conduct if they can appropriately be held to account for what they have done.<sup>3</sup> I shall furthermore assume that such holding to account paradigmatically includes expressing certain reactive attitudes toward them, such as resentment or indignation, and that it might additionally include imposing various forms of adverse treatment on them.<sup>4</sup> More precisely, I shall assume that if an agent is morally responsible (blameworthy) in this sense, then it will be fair (just, deserved) to react to their conduct in these ways.

Like many other theorists, I will moreover adopt a *moral competence requirement* for moral responsibility.<sup>5</sup> However, I will focus exclusively on the cognitive (or epistemic) component of this requirement (as opposed to the volitional one; see Talbert 2016: ch.4<sup>6</sup>). Roughly speaking, the cognitive moral competence requirement (henceforth CMCR) is concerned with an agent's ability *to recognize* that certain ways of behaving are morally wrong. Somewhat more precisely, we can spell out CMCR as follows:

---

<sup>2</sup> I will briefly discuss other (potential) varieties of moral responsibility, in particular, responsibility-as-attributability (Watson 1996; Shoemaker 2017) in section 5.2 of this paper.

<sup>3</sup> This helpful way of putting things comes from Randolph Clarke (2024: 62). I shall be exclusively concerned with narcissistic agents' *blameworthiness*. While the issue of narcissistic agents' praiseworthiness is an intriguing one, it is beyond the scope of this paper.

<sup>4</sup> Again, I (largely) follow Clarke (2024: 62–63) here.

<sup>5</sup> In his 'Stanford Encyclopedia' article on moral responsibility, Matthew Talbert (2024) lists the following theorists as proponents of (some form of) moral competence requirement: Benson (2001); Fischer & Ravizza (1998); Fricker (2010); Levy (2003); Russell (1995, 2004); Wallace (1996); Watson (1987); and Wolf (1987, 1990).

<sup>6</sup> Note that Talbert himself rejects any form of moral competence requirement. I shall elaborate on his position in section 5.2.

**CMCR:** Full moral responsibility requires the capacity to (by and large) form accurate moral assessments of the situations one confronts.<sup>7</sup>

Going forward, I shall assume that cognitive moral competence and impairments of that competence come in degrees (like many other competences and impairments thereof do). My theoretical rationale for this assumption is that it is highly plausible to assume that the ability which grounds cognitive moral competence—the ability to form accurate moral assessments of the situations one confronts—is something that also comes in degrees.<sup>8</sup>

A type of case that has motivated responsibility scholars to introduce CMCR is that of an agent who, due to their upbringing, ends up with a set of highly problematic values. This is illustrated by Susan Wolf's (1987) much-discussed fictional example of JoJo, which we can summarize as follows:

**JoJo:** 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' (Wolf 1987: 379). In consequence, JoJo comes to develop values that are very much like his father's. As an adult, JoJo does the same sort of atrocious things his father did, for example, sending people to prison, death, or torture chambers on the basis of a whim (see Wolf 1987: 379).

As Wolf furthermore emphasizes, though, JoJo's background story simultaneously seems to mitigate his responsibility or even fully exempt him. She writes:

In light of JoJo's heritage and upbringing—both of which he was powerless to control—it is dubious at best that he should be regarded as responsible for what he does. It is unclear whether anyone with a childhood such as his could have developed into anything but the twisted and perverse sort of person that he has become. (Wolf 1987: 380)

Importantly, as Talbert (2016: 103) notes, Wolf's idea is not that JoJo lacks (full) responsibility for his conduct because he had no control over his upbringing. Rather, it is that he is not responsible, or at least not fully so, because his upbringing (over which he had no control)

---

<sup>7</sup> The qualification 'by and large' is necessary because it seems quite implausible to assume that perfect moral competence—a feature that virtually no (human) agent possesses—is necessary for full moral responsibility. Plausibly, the capacity to form accurate moral assessments of the situations one confronts is itself grounded in a set of capacities, such as, for example, the capacity to correctly identify the morally relevant features of the situations one confronts, the capacity to grasp and weigh the moral reasons pertinent to these situations, and the capacity to apply moral principles to these situations.

<sup>8</sup> Many thanks to an anonymous reviewer for urging me to clarify this point.

'seems to have *damaged* him in a particular way' (Talbert 2016: 103, my emphasis), namely, by significantly impairing his ability to arrive at accurate moral assessments in a wide range of situations.

Problematic formative circumstances are not the only factor that has been claimed to (potentially) impair cognitive moral competence. According to some theorists (see, e.g., Wolf 1987; Benson 2001; Levy 2003), social or cultural environments can do so as well.

As several proponents of CMCR have convincingly argued, we need not assume that an agent is, strictly speaking, *unable* to recognize the wrongness of a certain behaviour for it to be true that this agent's responsibility is mitigated. Rather, it will suffice that it is *extremely difficult* for the agent to do so (see, e.g., Benson 2001; see also Levy 2003).

One key theoretical reason for accepting CMCR has to do with considerations of fairness. The 'argument from fairness' (see Wallace 1996; see also Watson 1996 and Levy 2009) can be summarized as follows:

**Argument from Fairness:** If A performs a morally wrong action X, but it is, due to an impairment in cognitive moral competence, extremely difficult for A to recognize that X is wrong, then it will be unfair to expose A to expressed resentment, expressed indignation, and/or various forms of adverse treatment on account of that conduct.<sup>9</sup> Or, at least, it will be, other things being equal, unfair to expose A to the same amount of such blaming responses as an agent who is not so impaired.<sup>10</sup>

The basic line of thinking contained in this argument is very intuitive and seems to provide us with a strong reason for adopting CMCR.

In the remainder of this paper, I shall provide an argument for the claim that narcissistic agents have impaired cognitive moral competence and therefore mitigated responsibility.<sup>11</sup> Note that to find this argument convincing, one need not agree with the claim that any of the specific factors mentioned above—problematic formative circumstances, social or cultural

---

<sup>9</sup> I shall remain neutral on the (contested) issue of whether unexpressed resentment and indignation can also be unfair.

<sup>10</sup> I am focusing on fairness, but it seems very plausible that the argument just offered could also be rephrased as an argument from *justice* or *desert*.

<sup>11</sup> More precisely, I will argue that narcissistic agents display various degrees of impairments of cognitive moral competence and that, therefore, some narcissistic agents (merely) have mitigated responsibility while others may be fully exempted from responsibility, depending on the severity of the impairments they display (on this point, see section 4.3).

environment—can impair an agent’s moral competence. This is because my argument focuses on a different factor—one that, roughly speaking, has to do with an agent’s basic personality structure.<sup>12</sup>

### 3. Narcissism: A Complex and Non-Homogeneous Condition

To make my case for the claim that narcissistic agents have impaired moral competence, it will be essential to gain a clear understanding of what narcissism *is*.

Following much of the recent empirical literature, I take the term ‘narcissism’ to refer to a personality structure or configuration of personality traits that agents can possess to different degrees (see, e.g., Krizan and Herlache 2017; Miller, Hoffman, *et. al.* 2017; Campbell and Christ 2020). On this picture, narcissism is distributed along a spectrum and should not be equated with narcissistic personality disorder. Narcissistic personality disorder is what we find at the high end of the narcissistic spectrum. But there will be many agents who qualify as narcissistic without having narcissistic personality disorder.

There is a growing consensus that there are two types of narcissism or, more precisely, narcissistic tendencies: *grandiose* and *vulnerable* narcissism (Miller, Hoffman, *et. al.* 2011; Krizan and Herlache 2017; Campbell and Christ 2020). Both share the personality trait *antagonism*. In a nutshell, someone with this trait will tend to mistrust and manipulate others, engage in egoistic behaviour, have difficulties complying with rules, and be immodest and callous (Campbell and Christ 2020). That said, the two types of narcissism combine antagonism with different features. In grandiose narcissism, antagonism is combined with the personality trait *extraversion* and high self-esteem. In vulnerable narcissism, antagonism is combined with the personality trait *neuroticism* and low self-esteem. These two types of

---

<sup>12</sup> Aleksandar Fatic also maintains that narcissism involves a form of ‘*moral incompetence*’ (2023a: 160). Despite similar terminology, Fatic’s take on the matter is very different from mine. For instance, Fatic argues that narcissistic agents’ ‘*moral incompetence*’ has its source in an ‘*emotional incompetence*’ (2023a: 159), specifically, a lack of empathy. As will become clear, I reject this notion and instead maintain that the source of narcissistic agents’ impaired moral competence is their psychological entitlement (see section 4). Moreover, Fatic does not think that narcissistic agents’ ‘*moral incompetence*’ mitigates their responsibility, since, according to him, narcissistic agents have the moral duty ‘to acquire the moral and emotional competence required for the making of moral choices’ (2023a: 165). For a helpful critical discussion, see Pies (2023).

narcissism are not mutually exclusive. An agent can display both grandiose and vulnerable narcissistic tendencies. Such a 'mixed type' can be found in many agents at the high end of the narcissistic spectrum (see Krizan and Herlache 2017; Campbell and Christ 2020).

Narcissistic agents are famous for behaving in unempathetic ways. In fact, the few extant philosophical discussions of narcissism and moral responsibility (accountability) tend to focus on this feature (Fatic 2023a; Fatic 2023b; Shoemaker 2024: ch. 5; see also Lamb 2022). I do not follow this approach for two reasons.

First, focusing on empathy deficits does not allow one to differentiate narcissism from other members of the so-called 'dark triad' (Paulhus and Williams 2002), such as psychopathy or Machiavellianism. This is because all dark triad constructs share this feature (see Krizan and Herlache 2017: 7–8).

Second, and more important, recent empirical evidence suggests that narcissistic agents' empathy deficits reflect more of a *lack of motivation* to empathize with others rather than a *lack of ability* (see, in particular, the meta-review by Urbonaviciute and Hepper 2020).<sup>13</sup> However, the fact that an agent behaves badly due to being insufficiently motivated to empathize with others is, in general, no reason to regard that agent as less than fully responsible for that behaviour.

Instead, I shall focus on *psychological entitlement*, a feature that has been largely overlooked by the few extant philosophical contributions on narcissism and responsibility.<sup>14</sup> According to one influential contribution, entitlement<sup>15</sup> refers to 'a global and pervasive sense that one deserves more and is entitled to more than others' (Campbell, Bonacci, *et. al.* 2004: 30–31).<sup>16</sup>

---

<sup>13</sup> Ronald Pies (2023) also stresses this point in his comment on Fatic (2023a).

<sup>14</sup> Shoemaker (2024, chs. 3 and 4) occasionally touches on this feature. However, it plays no role in his discussion of narcissistic agents' accountability. According to Shoemaker (2024, ch. 5, 69–70; ch. 7: 113) narcissistic agents are not accountable for their behaviour, due to their empathy deficits. However, he additionally maintains that they are an apt target of 'mocking-blame' which he claims to be 'a valuable, but previously unrecognized form of blame' (Shoemaker 2024, ch. 4: 58). Whether Shoemaker is correct about this is a topic for another paper.

<sup>15</sup> For ease of exposition, I shall usually omit the term 'psychological' and simply talk about entitlement.

<sup>16</sup> W. Keith Campbell and colleagues' (2004) concept of psychological entitlement is quite similar to Zlatan Krizan and Anne Herlache's concept of 'entitled self-importance', which is defined as 'a sense of oneself and one's needs being special and more important than others' (Krizan and Herlache 2017: 4).

In order to get a clearer grasp on this feature, it will be helpful to look at the scale by which it is measured (see Campbell, Bonacci, *et. al.* 2004). To assess their ‘entitlement score’, subjects are asked to indicate how well each of the following nine items reflects their own beliefs (using a 7-point scale, where 1 indicates strong disagreement and 7 indicates strong agreement):

1. I honestly feel I’m just more deserving than others.
2. Great things should come to me.
3. If I were on the Titanic, I would deserve to be on the *first* lifeboat!
4. I demand the best because I’m worth it.
5. I do not necessarily deserve special treatment.
6. I deserve more things in my life.
7. People like me deserve an extra break now and then.
8. Things should go my way.
9. I feel entitled to more of everything.<sup>17</sup>

The ‘entitlement scale’ measures a sense of special deservingness and special entitlement. Importantly, though, this sense is not of a local kind. That is, it does not refer to *specific* situations and is not based on *specific* merit (as in, for example, ‘I deserve a good grade for this test because I performed well on it’ or ‘I deserve a bigger piece of this cake because I baked it’). Rather, the scale measures a *global* sense of deservingness and entitlement (as reflected, for example, by the items ‘I honestly feel I’m just more deserving than others’; ‘Things should go my way’; ‘I feel entitled to more of everything’).

Entitlement has been found to be quite stable (Campbell, Bonacci, *et. al.* 2004; Krizan and Herlache 2017). Thus, narcissistic agents do not merely differ significantly from non-narcissistic agents in how entitled they are, but these differences also persist over time. Going forward, I shall assume that entitlement is one key component of narcissism (Campbell, Bonacci, *et. al.* 2004; see also Miller, Hoffman, *et. al.* 2011) and, perhaps, even its essential feature (Krizan and Herlache 2017).<sup>18</sup>

---

<sup>17</sup> See Campbell, Bonacci, *et. al.* (2004). Strong agreement with the items on the scale indicates *high* entitlement, except for Item 5, where the converse relationship holds.

<sup>18</sup> Krizan and Herlache (2017) argue that regarding entitlement as the essential feature of narcissism has two key theoretical advantages: (i) it allows one to explain what the two different types of narcissism—grandiose and vulnerable narcissism—have in common and (ii) simultaneously allows a clear delineation from other ‘dark triad’ constructs.

Joshua Miller and colleagues have conjectured that there may be different ‘rationale[s]’ (Miller, Hoffman, *et. al* 2011: 1032) for entitlement in grandiose and vulnerable narcissism. Agents with grandiose narcissistic tendencies, they maintain, ‘may believe they are entitled to special treatment because they are better than others (e.g., more attractive, more intelligent, more likable)’. By contrast, agents with vulnerable narcissistic tendencies ‘may believe they deserve special consideration because of their fragility’ (Miller, Hoffman, *et. al.* 2011: 1032), that is, because they conceive of themselves as being more sensitive or suffering more than others. Thus, there may be different rationales for psychological entitlement depending on whether an agent displays grandiose narcissistic tendencies, vulnerable narcissistic tendencies, or a combination of both. The important point, however, is that having narcissistic tendencies of either kind involves a global sense of special deservingness and entitlement.

Given psychological entitlement’s global character, it is unsurprising that this feature has been found to be linked ‘to a pattern of selfish and self-serving beliefs and behaviors’ in a wide variety of settings (Campbell, Bonacci, *et. al.* 2004: 43, see also Campbell, Bush, *et. al.* 2005).<sup>19</sup> That said, entitlement does not just systematically shift narcissistic agents’ behaviour in the direction of ‘self over others’. As I shall argue in the next section, it also impairs their moral competence.<sup>20</sup>

---

<sup>19</sup> Among other things, subjects high in entitlement were found to take more candy from a bowl of Halloween candy that they believed to be designated for children in the developmental laboratory (Campbell, Bonacci, *et. al.* 2004: 36–37). When asked to imagine that they worked for a company facing a cost-cutting situation, subjects high in entitlement claimed that they deserved to retain more salary than their colleagues and suggested greater salary cuts for their colleagues than for themselves (Campbell, Bonacci, *et. al.* 2004: 37). Moreover, subjects high in entitlement were found to report ‘being essentially more selfish in their romantic relationships’ (Campbell, Bonacci, *et. al.* 2004: 39)

<sup>20</sup> After revising the paper, I discovered that the conception of ‘psychological entitlement’ put forward by Campbell and colleagues (2004) largely overlaps with the (apparently) armchair conception of ‘moral entitlement’ developed by Aaron James (2012) in his popular book on ‘assholes’. According to James (2012, ch. 1), ‘moral entitlement’ is a defining feature of being an asshole—a category which he claims to be both narrower and wider than that of the narcissist (see James, 2012: 21, n. 16 and 36, n. 36). Due to the significant overlap in the underlying conceptions of entitlement, there are also some similarities between the claims made in the next section of my paper and the claims James puts forward in his astute discussion of the ‘asshole’ (which perhaps might be better characterized as mid-level grandiose narcissist). However, there are also some crucial differences, both in the specific way entitlement is supposed to operate and in the normative consequences drawn. Concerning the asshole’s responsibility, James (2012, ch. 4) seems to advocate a position similar to the one that I shall reject in section 5.2.

## 4. Narcissism as Impaired Moral Competence

### 4.1 The Global Normative Specialness Illusion and Narcissistic Distribution Bias

If psychological entitlement is both pervasive and stable (see section 3), then it seems to follow that an agent with this feature will not merely believe in some *specific* situation that they are more deserving and (hence) entitled to more than others, but will rather believe this *across the board* (at least at some level)—and will continue to believe this over time. However, in view of this, narcissistic agents can be described as being subject to a sort of stable and pervasive normative illusion, which I shall call the *global normative specialness illusion*. The global normative specialness illusion encompasses the deep-seated, stable, and illusionary beliefs that one is

(1) *generally* more deserving than others (*global special deservingness*)

and (thus)

(2) *generally* entitled to more than others (*global special entitlement*).

It seems plausible to assume that an agent who is subject to the global normative specialness illusion will systematically misrepresent their own deservingness and what they are entitled to. More specifically, being subject to the global normative specialness illusion impairs an agent's cognitive moral competence, or so I shall argue. For ease of exposition and because it may be conjectured to be the more basic stance of the two, I shall focus on *global special deservingness*. That said, my argument can simply be carried over to the case of global special entitlement.

The proposal I want to put forward, more precisely, maintains that global special deservingness influences and distorts the way a narcissistic agent spontaneously assesses a given situation and its morally relevant features. In maintaining that global special deservingness distorts a narcissistic agent's *spontaneous* moral assessments, I want to suggest that global special deservingness operates in a way analogous to a *bias*. To clarify, this suggestion does not rule out that narcissistic agents may also consciously and reflectively endorse the belief that they are generally more deserving than other agents. Rather, the idea

is that global special deservingness will exert its distorting influence independently of this, namely, by directly and automatically influencing the way a given situation is ‘perceived’.<sup>21</sup>

More specifically, we should expect (at least) two types of systematic moral distortions given global special deservingness, namely, (i) overestimating the share of some (desired) good that is due to oneself and (ii) underestimating the share of some (desired) good that is due to others. In view of this, global special deservingness can be characterized as either involving or leading to a *distribution bias*: because they have the deep-seated illusionary belief that they are generally more deserving than others, it will systematically *seem* to narcissistic agents that they deserve a larger share of some (desired) good than they really do and, conversely, that other agents deserve a smaller share of this good than these other agents really do. Or, in other words, due to displaying global special deservingness, a narcissistic agent will systematically perceive the distribution of some (desired) good as being

- (i) ‘only what is due to me’ even if the distribution is actually very unfair or else as
- (ii) ‘not at all what is due to me’ even if the distribution is actually perfectly fair.

To this point, the proposed model has left open which goods fall within the scope of narcissistic distribution bias. In principle, many different goods could be filled in here. Some obvious candidates are classic ‘status goods’ (for example, money, promotions, awards, and fame).

This is illustrated by the following two examples:

**Immense Bonus:** Beth is the CEO of a mid-sized company. The year is coming to an end and Beth is considering whether to issue bonus payments. The company is struggling financially, mainly as a result of poor decisions that Beth has made. The fact that the company has not yet gone bankrupt is largely due to the tireless efforts of Beth’s assistant, Ann. So, if anyone deserves a bonus, it is, in fact, Ann. But this is not how things seem to Beth who, in virtue of being narcissistic, displays global special deservingness. Rather, it seems to Beth that she herself indisputably deserves a large bonus—much more so than Ann does. When Beth decides to make a very large payment to herself and none to Ann, it appears to Beth that she (Beth)

---

<sup>21</sup> In a similar vein, Chandra Sripada (2022) has recently stressed the role of automatically produced systematically inaccurate evaluations for explaining addiction and various other psychiatric conditions.

is only receiving *what it is due to her* and that things are precisely as they *should* be.

**Stolen Data:** Ann and Beth work in the same lab. With great diligence and a significant time investment, Ann has managed to produce a set of very valuable data. If Ann got a publication out of this in a top journal, it would be highly deserved. But this is not how things seem to Beth, who, in virtue of being narcissistic, displays global special deservingness. Rather, it seems to Beth that she deserves to publish in a top journal much more than Ann does and that this is actually *long overdue*. Having access to Ann's data set, Beth quickly writes up an article based on it and submits it as her own work.

In these examples, it *appears* to the narcissistic agent (Beth) to be the case that she deserves a certain desired 'status good' (a large bonus payment, a publication in a highly-ranked journal) much more than other agents do (even though she actually does not). She consequently brings about a distribution of that good that *seems* to her to be *only what is due to her* (even though that distribution is actually utterly unfair).

Assuming that Beth has grandiose narcissistic tendencies, her distorted spontaneous 'desert beliefs' might well be accompanied by various distorted descriptive beliefs about her own abilities or performance (see section 3). She might, for instance, (wrongly) believe that she is a highly competent CEO (Immense Bonus), or a brilliant, yet unrecognized, scientist, while Ann is merely a mediocre one (Stolen Data). Beth might then use these distorted descriptive beliefs to rationalize her distorted spontaneous desert beliefs. For instance, Beth might construe the immense bonus she pays to herself as a *just reward for her exceptional leadership* or the theft of the data as a *just redistribution of scientific fame* (the scientific fame now goes where it is actually deserved).

Note that, on my proposed model, the causal history is the reverse of what it usually is. Normally, our desert beliefs are based on our beliefs about the specific things we did (or other descriptive beliefs). Ann, for instance, might believe that she has rescued the company from bankruptcy (or produced a set of very valuable data). Based on these beliefs, she might then form the belief that she deserves a large bonus payment (or a publication in a top journal). Beth, by contrast, due to displaying global special deservingness, spontaneously forms the

belief that she deserves a large bonus payment (or a publication in a top journal), and then uses her (equally distorted) descriptive beliefs in order to rationalize her desert beliefs.<sup>22</sup>

Importantly, the model just presented can also account for vulnerable narcissistic tendencies. Here, narcissistic distribution bias will characteristically manifest itself with respect to (what might be called) 'relationship goods', such as, for example, others' attention, time, or emotional support.<sup>23</sup> This can be illustrated by the following two examples:

**Work Incident:** Beth has had (what in fact amounts to) a very minor incident at work (a colleague interrupted her while she was giving a presentation). She decides to call her friend Ann about it. Ann tells her that she does not have time for a conversation right now because she must prepare for an important meeting, but offers to talk to Beth about the incident in the evening. In the past, Ann has often listened to Beth complain about her problems, even when it has interfered with Ann's own projects. Ann's reaction, therefore, appears perfectly reasonable. But this is not how things seem to Beth, who, in virtue of being narcissistic, displays global special deservingness. Instead, it seems to Beth that Ann is not giving her the time she deserves to talk about a very serious incident and that Ann should stop preparing for her meeting and listen to her right away. Feeling very angry about Ann's refusal, Beth abruptly hangs up on her and does not return her texts for a fortnight.

**Wedding Blues:** Beth is attending her godchild Ann's wedding. The ambiance at the wedding reception is cheerful and everyone seems to be having a great time. However, with the focus of attention being clearly on the bridal couple, Beth grows increasingly uncomfortable. While this unequal distribution of attention seems natural to everybody else, it does not seem so to Beth, who, in virtue of being narcissistic, displays global special deservingness. Rather, it seems to Beth that there is a lot of unnecessary fuss around the bridal couple, that she herself is not receiving the attention that is due to her, and that things are not at all as they should be. She, consequently, becomes very sullen and monosyllabic. When another guest gives a very thoughtful speech, her facial expression grows increasingly sombre. Without excusing herself or saying goodbye to Ann or anyone else, Beth gets up and leaves the reception.

In these examples, it *appears* to the narcissistic agent (Beth) to be the case that she deserves a much bigger share of certain 'relationship goods' (others' attention, time, emotional support)

---

<sup>22</sup> My claim that in narcissism, distorted desert beliefs come first and then may be rationalized by distorted descriptive beliefs is, admittedly, a conjecture, but a plausible one. It enables us to best account for the fact that the entitlement displayed by narcissistic agents is 'intrapsychically global and pervasive' (Campbell, Bonacci, *et. al.* 2004: 31) and to capture the specific predicament of narcissistic agents.

<sup>23</sup> To clarify, my claim here is that the distribution bias displayed by agents with vulnerable narcissistic tendencies will characteristically include relationship goods (for empirical support of this claim, see Krizan and Herlache 2017: 13). My claim is not that it is restricted to relationship goods.

than she actually does. Consequently, others' disinclination to provide her the desired amount of these goods strikes her as being *not at all what is due to her*, leading to the impression that she is being wronged when in fact she is not.

Assuming that Beth has vulnerable narcissistic tendencies in the latter cases, her distorted spontaneous desert beliefs might well be accompanied by various distorted descriptive beliefs about herself and her situation (see section 3). For instance, Beth might believe that her colleague's interruption constituted a *major* slight and has caused her an *exceptional* amount of distress (Work Incident), or that Ann and the other guests are practically ignoring her even though her own life is *exceptionally difficult* right now (Wedding Blues). She might then use these distorted descriptive beliefs to rationalize her distorted desert beliefs. Beth might, for instance, construe Ann's behaviour as selfish (Work Incident) or inconsiderate (Wedding Blues). And she might construe her own passive-aggressive behaviour as an instance of *just punishment*, that is, as something Ann (and the other guests) deserve for refusing to give her the emotional support, consideration, and/or attention which she (Beth) deserves. Note that, on the proposed model, the causal history is, again, the reverse of what it usually is. Due to displaying global special deservingness, Beth spontaneously believes that she deserves Ann's immediate emotional support (or a larger share of the attention than she is getting), and then uses her (equally distorted) descriptive beliefs about her situation to rationalize her distorted desert beliefs.

#### **4.2 Narcissistic Distribution Bias and Self-Correction**

In the preceding section, I argued for the following claim: the fact that narcissistic agents display global special deservingness leads them to form numerous spontaneous distorted beliefs about how a wide variety of goods (bonus payments, publications in top journals, emotional support, attention, etc.) would have to be distributed so that people in general and they themselves in particular would get what they deserve. I have furthermore argued that narcissistic agents may use various distorted descriptive beliefs that are rooted in their grandiosity and vulnerability, respectively, to rationalize their distorted desert beliefs and the further distorted moral beliefs which result from them.

At this point, a natural objection arises. Narcissistic agents may not be responsible for 'finding themselves' with distorted moral assessments of the situations they confront. They are, however, fully responsible for not correcting these distortions, especially if others have pointed them out to them.

However, this objection does not withstand closer scrutiny. It is another key feature of the entitlement we find in narcissism and the distribution bias it entails that it undermines the correction of the distortions it produces. There are several reasons for this.

First, entitlement is correlated with a tendency to react aggressively to criticism. For instance, Campbell and colleagues (2004) found that entitled agents who received a poor evaluation for an essay 'lashed out and behaved aggressively against the person who criticized them' (2004: 41). There seems to be little reason to expect that entitled agents' reactions to *moral* criticism will be any different.

Second, many instances of moral criticism will simply seem illegitimate to an entitled mind, especially if they demand equal consideration, rights, or obligations for all (as many instances of moral criticism do). Take a case of criticizing someone for ignoring rules that apply to all (for example, skipping a queue, making phone calls when traveling in the 'quiet zone', failing to wear a mask when visiting a hospital, etc.). The denial of special treatment implicit in such criticism must (at some level) seem illegitimate to an entitled agent. After all, such an agent believes 'deep down' that they very much *deserve* special treatment and (consequently) *have the right* to display these behaviours. But even non-entitled agents will typically not reform their behaviour in response to instances of moral criticism that appear illegitimate to them.

Third, moral criticism may itself be among the (negative) goods that fall in the scope of narcissistic distribution bias. Receiving moral criticism is often unpleasant. In view of this, entitled agents who are confronted with moral criticism might find themselves with the belief that *they simply do not deserve to be treated in such a disagreeable manner*, or, relatedly, that *they should receive (much) more praise and (much) less criticism* (than they actually do).<sup>24</sup>

---

<sup>24</sup> To illustrate, consider the following statement by Donald Trump about the Republican candidates for the 2022 Midterm elections: 'Well, I think if they win, I should get all the credit, and if they lose, I should

Plausibly, the three considerations just mentioned significantly impair narcissistic agents' ability to deal constructively with moral criticism and hence their ability to correct their distorted spontaneous moral assessments of the situations they confront. Hence, the very same deep-seated psychological feature that makes a narcissistic agent *more* likely to form a distorted moral assessment of a given situation will simultaneously make it *less* likely that they will be able to use critical feedback from others to correct that assessment and form more accurate judgments in the future.

One may wonder whether entitlement and the distribution bias it entails are unique in this respect. One could ask, for example, if racial or sexist biases do not undermine an agent's ability to use critical feedback to self-correct in a similar way.<sup>25</sup> Adequately discussing this question would require a paper of its own, but let me offer a brief preliminary reply. Arguably, racial or sexist biases significantly impair an agent's ability to deal constructively with instances of moral criticism issued by members of the 'targeted' group (that is, the group toward which the agent is biased). Entitlement and the distribution bias it entails, by contrast, significantly impair an agent's ability to deal constructively with instances of moral criticism irrespective of who is the source of that criticism. This is because the three considerations listed above provide serious *and general* obstacles to using critical social feedback to self-correct. For this reason, a narcissistic agent's capacity to use critical social feedback for self-correction may be regarded as more pervasively impaired than that of the (average) racist or sexist.<sup>26</sup>

---

not be blamed at all' (Rolling Stone, November 8, 2022, <https://www.rollingstone.com/politics/politics-news/trump-credit-gop-wins-midterms-no-blame-lose1234627288/>).

<sup>25</sup> Many thanks to an anonymous reviewer for raising this important question.

<sup>26</sup> The preceding does not rule out that there may be other features, either in the psychology or in the environment of a racist or sexist that may make self-correction very difficult (on environmental factors, see, e.g., Wolf 1987; Wolf 1990; Levy 2003). It is thus not surprising that some proponents of the moral competence requirement have maintained that some racists or sexists (likely) suffer from 'local' impairments of cognitive moral competence, and hence have mitigated responsibility with respect to certain contexts or situations (see, e.g., Wolf 1987; Wolf 1990: 121–22; Levy 2003). Other responsibility scholars have tried to mitigate the intuitive unpleasantness of this conclusion by arguing that the relevant agents are still fully responsible for their bad behaviour in some other important sense of the term 'responsible', for example, the attributability sense (see Nelkin 2015: 383, 386–87; on responsibility-as-attributability, see also n. 38).

### 4.3 Putting it all together: Narcissism as De Se Impaired Moral Competence

So far, I have argued for two main claims:

In virtue of displaying global special deservingness, narcissistic agents

- (1) form numerous spontaneous distorted beliefs about how a wide variety of goods would have to be distributed so that they themselves would get what they deserve.
- (2) have significant impairments in their ability to use feedback from others to correct their distorted spontaneous desert beliefs and the further distorted moral beliefs which result from them.

In virtue of having the effects just described, global special deservingness arguably impairs narcissistic agents' ability to form accurate moral assessments of the situations they confront and hence their cognitive moral competence. Due to displaying global special deservingness, it is extremely difficult for narcissistic agents '[to] normatively recognize and appreciate the world for what it is' (Wolf 1987: 383),<sup>27</sup> namely, a place where they are *not* generally more deserving than other agents and (hence) *not* generally entitled to a bigger share of positive goods (or a smaller share of negative goods). Consequently, their ability to arrive at accurate moral assessments of the situations they confront will be significantly impaired.

The impairment of cognitive moral competence we find in narcissism will, plausibly, be a gradualist one. As pointed out earlier (see section 3), narcissism is distributed along a spectrum and entitlement, too, comes in degrees. However, an agent who is strongly entitled will display global special deservingness to a significantly higher degree than an agent who is only mildly entitled. Consequently, the distorting effects on their moral assessments will be (much) stronger and their cognitive moral competence will be (much) more impaired compared to a mildly entitled agent.

However, given that impaired cognitive moral competence reduces moral responsibility (see section 2), it follows from the preceding that narcissism is a condition that mitigates moral responsibility and that the responsibility-mitigating effect will be stronger the higher an agent is located on the narcissistic spectrum (due to the fact that moving up the narcissistic spectrum

---

<sup>27</sup> To clarify, this quotation is taken from Wolf's discussion of JoJo (see section 2), i.e., Wolf is not discussing narcissism here.

will go along with an *increase* in entitlement and hence with a *decrease* in cognitive moral competence). Or, in other words, we are led to the conclusion that mildly narcissistic agents have somewhat mitigated responsibility, whereas highly narcissistic agents have significantly mitigated responsibility or may even be fully exempted from responsibility.

Before turning to objections, let me point out another interesting feature of impaired cognitive moral competence in narcissism. I have maintained that the entitlement displayed by narcissistic agents will impair their ability to form accurate moral judgments *in a wide range of situations*. In view of this, it may seem that the normative distortions we find in narcissism will be much more ‘global’ (Talbert 2016: 101) than what we see in (many) cases of ‘culturally induced’ impaired moral competence (see, e.g., Wolf 1987; Benson 2001; Levy 2003). That said, there simultaneously seems to be a clear *delimitation* to a narcissistic agent’s impaired cognitive moral competence. This is because it seems plausible to expect that global special deservingness will primarily impair a narcissistic agent’s ability to arrive at accurate moral assessments where they themselves are concerned. What I have argued for in this section is thus perfectly compatible with the assumption that a narcissistic agent is able to make accurate moral judgments about, for example, the obligations of inhabitants of rich countries to fight global poverty or climate change (while the accuracy of their moral judgments about the specific moral obligations that follow from this for themselves will, again, be hindered by global special deservingness). This explains why it is perfectly possible to imagine a highly narcissistic agent who is in charge of an NGO or (even more to the point) a highly narcissistic moral philosopher. The impairment of moral competence in narcissism can thus be accurately characterized as a *de se*-impairment.<sup>28</sup> It is something that primarily distorts a narcissistic agent’s moral assessments where *they themselves* are concerned.<sup>29</sup>

## 5. Objections

### 5.1 Responsibility for Impaired Cognitive Moral Competence

---

<sup>28</sup> Many thanks to Peter Schulte for suggesting this very helpful terminology to me.

<sup>29</sup> James (2012: 110) makes a similar observation about (what he calls) ‘the asshole’ (on this point, see also n. 20 of this paper).

One important objection against the reasoning offered in the preceding sections is as follows. Plausibly, if an agent is responsible for the feature that impairs their moral competence, then they will be fully responsible for their wrongful conduct. But narcissistic agents may well be responsible for their entitlement and, more specifically, for the beliefs which are at its core, namely, that they are generally more deserving and entitled to more than others.<sup>30</sup>

I am happy to concede that if an agent is responsible for their impaired cognitive moral competence, then their responsibility for the wrongful conduct they display is not mitigated. However, it seems questionable that this is the case when it comes to narcissism. More specifically, there are at least two serious problems with the view that narcissistic agents are responsible for the beliefs which are at the core of their entitlement.

First, remember that I rely on a notion of responsibility as accountability (see section 2). However, it is quite controversial that one can be accountable for having certain beliefs. Second, it is empirically plausible that narcissistic agents acquire the relevant beliefs at a point in their life, where due to reasons of immaturity, they still lack a sufficiently robust cognitive moral competence.<sup>31</sup> But given that I assume that cognitive moral competence is necessary for moral responsibility (see section 2), it follows that narcissistic agents cannot be responsible for acquiring the beliefs which are at the core of their entitlement. And once these beliefs are acquired, they are, plausibly, self-stabilizing for the reasons offered in the preceding (that is due to the fact that entitlement significantly impairs narcissistic agents' ability to self-correct; see section 4.2).

The considerations just listed arguably undermine the initial plausibility of the view that narcissistic agents are responsible for having the beliefs that they are generally more deserving and entitled to more than others, and, consequently, for their impaired cognitive moral competence. In fact, I contend that the above considerations make it legitimate to shift the burden of proof to proponents of this view.

---

<sup>30</sup> Many thanks to an anonymous reviewer for bringing this objection to my attention.

<sup>31</sup> There is substantial empirical evidence that narcissism, and the entitlement it involves, emerges as early as adolescence (if not earlier) and is 'fully there' in young adulthood (see, e.g., Carlson and Gjerde 2009; Cramer 2011; Orth, Krauss, and Back 2024).

## 5.2 Narcissism, Attributionism, and the Significance of Narcissistic Wrongdoing

The entitlement we find in narcissism does not merely impair narcissistic agents' cognitive moral competence. It also makes it true that their actions will often reflect objectionable evaluative judgments (or other objectionable attitudes). For instance, Beth's stealing the data might express the objectionable judgment that the fact that it will hurt Ann's career prospects and feelings if she takes Ann's data and publishes them as her own is not a sufficient reason to refrain from doing so. However, according to some responsibility scholars, this is *enough* for an agent to be fully morally responsible for that behaviour (e.g., Scanlon 1998; Talbert 2014; Hieronymi 2014; Smith 2015; for an overview, see Talbert 2022).<sup>32</sup> This approach has been labelled 'attributionism' (Talbert 2022) since it is fundamentally concerned 'with the morally significant features of an agent's orientation toward others that are attributable to her' (Talbert 2022: 44).<sup>33</sup>

Unlike proponents of attributionism, I believe that being fully responsible for one's wrongful behaviour requires more than that an objectionable evaluative judgment (or other objectionable attitude) is revealed through one's conduct, at least on the assumption that responsibility is understood as accountability (see section 2).<sup>34</sup> It also requires moral competence (for the reasons stated in section 2). Unfortunately, settling this dispute by comprehensively reviewing the arguments for and against a moral competence requirement will not be possible within the scope of this paper. So instead, I shall try to lessen one important concern one may have about adopting such a requirement, which has been articulated by Matthew Talbert, one of the most fervent proponents of attributionism.<sup>35</sup> This is the concern

---

<sup>32</sup> I follow Talbert (2022) in classifying the above theorists together.

<sup>33</sup> Attributionism needs to be distinguished from responsibility-in-the-attributability sense (Watson 1996). I will elaborate on this distinction below (n. 38).

<sup>34</sup> Proponents of attributionism, too, rely on a notion of responsibility as accountability, but additionally maintain, as Matthew Talbert has put it, that '*attributability is enough for accountability*' (Talbert 2022: 57). Hence, at first glance at least, their position seems to be in direct conflict with the view defended in this paper (see, however, n. 35).

<sup>35</sup> On close inspection, it becomes somewhat less clear, though, how big the conflict between the view defended in this paper and attributionism really is. For instance, in reading Talbert, one sometimes gets the impression that (some) proponents of attributionism are more concerned with *aptness* conditions for *judgments of blameworthiness* than with *fairness* conditions for *expressions of blame* (see, e.g., Talbert 2022: 65–66). If this impression can be corroborated, then the apparent conflict between attributionism and my view may, in the end, turn out to be an instance of talking past each other.

that, in accepting a moral competence requirement, ‘we risk losing our grip on the moral status and significance’ (Talbert 2016: 111) of the problematic treatment we or third parties receive from morally incompetent agents.

There is some initial intuitiveness to this line of reasoning. Suppose I am correct in claiming that, due to their impaired cognitive moral competence, it is unfair (unjust, undeserved) to react to narcissistic agents’ problematic behaviour with expressed resentment, expressed indignation and adverse treatment, or, at least, to display these reactions to the same degree as we would, other things being equal, toward a non-narcissistic agent (see section 2).<sup>36</sup> Suppose further that we, therefore, decide to refrain from displaying these reactions (or at least to attempt to do so). One might worry that this leads us to conceive of problematic behaviour displayed by narcissistic agents as ‘morally analogous to an involuntarily, accidentally, or justifiably imposed harm’ (Talbert 2016: 111) and, thereby, to fail to do justice to the victim’s perspective (see Talbert 2016: 111–113).

Let me offer two points in reply. First, note that there is also some oddity in the reasoning just offered. As Jeanette Kennett has pointed out, ‘the allocation of responsibility traditionally rests on some feature of the offender, not the victim’ (Kennett 2019: 152). But Talbert’s reasoning seems to reverse this traditional approach (as Kennett 2019 also notes).

Second, there are alternative ways we can keep our grip on the moral status and significance of narcissistic agents’ bad behaviour which do not involve reacting to their conduct with the kind of robust blaming responses I am concerned with in this paper. Among other things, we can still morally assess their behaviour and clearly label it as the morally unjustified behaviour it is.<sup>37</sup> One way of doing so would be to engage in *aretaic* assessments of narcissistic agents’ bad conduct, that is to assess it as, for example, egocentric, inconsiderate, unfair, disrespectful, or even cruel.<sup>38</sup> By engaging in these kinds of aretaic assessments, we

---

<sup>36</sup> Of course, the gradualistic nature of narcissistic agents’ impaired moral competence would also have to be taken into account (see section 4.3). I omit this complication here for ease of exposition.

<sup>37</sup> The following point is inspired by Kennett’s (2019) reply to Talbert’s (2014) discussion of the moral responsibility of psychopathic agents.

<sup>38</sup> Some responsibility scholars—most famously, Gary Watson (1996)—have maintained that engaging in aretaic assessments is itself a way of holding an agent morally responsible. Watson calls this the

can not only keep a firm grip on the moral status and significance of the bad behaviour displayed by narcissistic agents, but also conceive of it as behaviour that is very much ‘in character’ and thus far from accidental.

### 5.3 Narcissism, Harm, and Self-Protection

Some normative discomfort with my account might still remain in the form of what we can call the ‘harm’ objection. This objection goes as follows: My account maintains that an increase in entitlement will lead to a decrease in moral competence and hence moral responsibility (understood as accountability). It thus implies that highly narcissistic agents, in virtue of their being highly entitled, have significantly mitigated responsibility or may even be fully exempted from responsibility. But this conclusion, the objection continues, is simply too hard to bear. After all, such agents are likely to do significant harm to others.

My account implies that it will be unfair (unjust, undeserved) to react to highly narcissistic agents’ bad behaviour with robust blaming responses, such as expressed resentment, expressed indignation, and various forms of adverse treatment (see section 2). However, it does *not* imply that we must simply accept the damage done by these agents. We may still take measures to protect ourselves against the destructive effects of their entitlement—measures which can be justified by appealing to our right to self-defence.<sup>39</sup>

Spelling out in detail the concrete measures that can be justified by this right in the case of narcissism is beyond the scope of this paper. Let me, nonetheless, briefly hint at what some of these measures might look like. On the personal level, the right to self-defence arguably

---

‘attributability face’ of responsibility and distinguishes it from the ‘accountability face’ of responsibility. This view would seem to imply that there is *a sense* in which narcissistic agents have unmitigated responsibility, namely, responsibility-as-attributability. However, other theorists—e.g., Jeanette Kennett (2019)—have argued against the view that engaging in aretaic assessments is a way of allocating moral responsibility or, in other words, that there is an attributability face of responsibility. Unfortunately settling this dispute is not possible within the scope of this paper (nor is settling the dispute of whether there is, additionally, an ‘answerability face’ of responsibility, see, e.g., Shoemaker 2017). Luckily, doing so is not necessary to reject Talbert’s concern since the above reply works independently of whether or not we classify aretaic assessments as *responsibility* responses.

<sup>39</sup> Various theorists have argued—and, in my view, convincingly so—that it can be justified to harm others’ interests in self-defence even if they are not morally responsible for the morally problematic behaviour they display (see, e.g., Ellis 2003; Kennett and Fine 2004; Pereboom 2014, 2021). This basic line of thought might also provide a promising starting point for building a theory of *criminal* responsibility of narcissistic agents’ wrongdoing (which is beyond the scope of this paper).

implies the right to distance oneself from a narcissistic agent or even to end the relationship completely (depending on the severity of the bad behaviour displayed). Importantly, this right could very well trump any special obligations we have toward a narcissistic agent (such as obligations towards family members or friends).

On the societal level, the right to self-defence may justify making important institutions 'narcissism-proof'.<sup>40</sup> This could involve taking steps to prevent narcissistic agents from accumulating and abusing power (for example, by increasing checks and balances or by introducing rotation systems and shared responsibilities) and it could also involve taking steps to reform leader selection (see also Campbell and Christ 2020, ch. 8).

Concerning the latter suggestion, one could imagine adopting leader-selection methods that do not favour selecting narcissistic agents or that actively prevent highly narcissistic agents from accessing certain leadership positions. This somewhat radical measure clearly demands a more thorough defence than I can offer here. But let me finish this section by offering some preliminary support for it, which comes in the form of an analogy: Most would agree that, for the safety of all, we ought not allow severely vision-impaired persons to become pilots, even if these persons are not responsible for being severely vision-impaired. Similarly, it could be argued that we ought not allow highly narcissistic agents to occupy important leadership positions, even if these agents are not responsible for being highly narcissistic. This is because their significantly distorted 'normative vision' renders it very likely that once they have arrived in these positions, they will behave in ways which, while aggressively promoting their own self-interest, will hurt the common good.

## **6. Conclusion**

I have defended the view that narcissistic agents have mitigated responsibility (accountability) or, in some cases, may even be fully exempted because they have impaired cognitive moral competence. More specifically, I have argued that narcissistic agents' entitlement will cause

---

<sup>40</sup> To clarify, the following measures are meant as preliminary and debatable suggestions for how to protect society against the significant harm done by highly narcissistic agents. Assessing whether these suggestions withstand closer scrutiny would require a paper of its own.

them to form distorted moral assessments in a wide range of situations, while simultaneously impairing their ability to use feedback from others to correct those distortions.

Given that narcissistic agents' entitlement disposes them to act in ways that can be characterized as, for example, egocentric, inconsiderate, disrespectful, or unjust, it seems clear that its effects will often be quite negative. Importantly, though, the view I have defended in this paper does *not* entail that we must simply accept the damage that narcissistic agents do. Rather, it is fully compatible with the view that we may take various steps to protect ourselves against the destructive effects of their entitlement.<sup>41</sup>

## 7. References

- Benson, Paul H. (2001) 'Culture and responsibility. A reply to Moody-Adams', *Journal of Social Philosophy* **32**: 610–620.
- Campbell, Keith W., Bonacci, Angelica, M., Shelton, Jeremy, Exline, Julie J. and Brad J. Bushman (2004) 'Psychological Entitlement. Interpersonal Consequences and Validation of a Self-Report Measure', *Journal of Personality Assessment* **83**: 29–45. doi: [10.1207/s15327752jpa8301\\_04](https://doi.org/10.1207/s15327752jpa8301_04).
- Campbell, Keith W., Pierce Bush, Carrie, Brunell, Amy B. and Jeremy Shelton (2005) 'Understanding the Social Costs of Narcissism. The Case of the Tragedy of the Commons' *Pers Soc Psychol Bull* **31**: 1358–1368. doi: [10.1177/0146167205274855](https://doi.org/10.1177/0146167205274855).
- Campbell, Keith W. and Carolyn Christ (2020) *The New Science of Narcissism. Understanding One of the Greatest Psychological Challenges of Our Time—and What You Can Do About it*, Boulder. Sounds True.
- Carson, Kevin S. and Per F. Gjerde (2009) 'Preschool personality antecedents of narcissism in adolescence and young adulthood: A 20-year longitudinal study', *Journal of Research in Personality* **43**: 570–578. doi: [10.1016/j.jrp.2009.03.003](https://doi.org/10.1016/j.jrp.2009.03.003).
- Cramer, Phebe (2011) 'Young adult narcissism: A 20 year longitudinal study of the contribution of parenting styles, preschool precursors of narcissism, and denial', *Journal of Research in Personality* **45**: 19–28. doi: [10.1016/j.jrp.2010.11.004](https://doi.org/10.1016/j.jrp.2010.11.004).

---

<sup>41</sup> The paper was accepted for publication by the Australasian Journal of Philosophy on December 24, 2024. Special thanks go to Keith W. Campbell for invaluable scientific advice, to Leonhard Menges and Chandra Sripada for repeated rounds of very insightful written feedback, to Peter Schulte for extremely helpful discussions, as well as to two reviewers from the Australasian Journal of Philosophy for engaging deeply with the paper and for providing two sets of very astute comments which greatly improved the manuscript. I would also like to thank Alida Mühlbauer, Dorothea Debus, Sonja Dembíc, David Heering, Erasmus Mayr, Paul Russell, Jay Wallace, and Shawn Wang for further very helpful comments and/or discussion. Previous versions of the paper were presented in the advanced seminar of the Salzburg Ethics Group (Salzburg, March 2024), at the workshop 'Responsibility in a Strawsonian Key' (Erlangen, June 2024) and the 31th Meeting of the European Society for Philosophy and Psychology (Grenoble, July 2024). I would like to thank the audiences at these events for further helpful feedback. I would also like to thank Lex Academic for proofreading an earlier version of this paper, and Claire Davis for proofreading the final version.

- Clarke, Randolph (2024) 'Desert of Blame', *Philosophy and Phenomenological Research* **108**: 62–80. doi: [10.1111/phpr.12952](https://doi.org/10.1111/phpr.12952).
- Ellis, Anthony (2003) 'A Deterrence Theory of Punishment', *Philosophical Quarterly* **53**: 338–351.
- Fatic, Aleksandar (2023a) 'Narcissism as a Moral Incompetence', *Philosophy, Psychiatry, and Psychology* **30**: 159–167. doi: [10.1353/ppp.2023.a899946](https://doi.org/10.1353/ppp.2023.a899946).
- Fatic, Aleksandar (2023b) 'Why Narcissists are Morally Responsible', *Philosophy, Psychiatry, and Psychology* **30**: 177–180. doi: [10.1353/ppp.2023.a899949](https://doi.org/10.1353/ppp.2023.a899949).
- Fischer, John Martin and Mark Ravizza (1998) *Responsibility and Control. A Theory of Moral Responsibility*. Cambridge University Press.
- Fine, Cordelia and Jeanette Kennett (2004) 'Mental Impairment, Moral Understanding and Criminal Responsibility: Psychopathy and the Purposes of Punishment', *International Journal of Law and Psychiatry* **27**: 425–443. doi: [10.1016/j.ijlp.2004.06.005](https://doi.org/10.1016/j.ijlp.2004.06.005).
- Fricker, Miranda (2010) 'The Relativism of Blame and Williams's Relativism of Distance', *Aristotelian Society Supplementary* **84**: 151–177. doi: [10.1111/j.1467-8349.2010.00190x](https://doi.org/10.1111/j.1467-8349.2010.00190x).
- Hieronymi, Pamela (2014) 'Reflection and Responsibility: Reflection and Responsibility', *Philosophy & Public Affairs* **42**: 3–41. doi: [10.1111/papa.12024](https://doi.org/10.1111/papa.12024).
- James, Aaron (2012). *Assholes. A Theory*. Nicholas Brealey Publishing.
- Jaworska, Agnieszka (2007) 'Caring and Internality', *Philosophy and Phenomenological Research* **74**: 529–568. doi: [10.1111/j.1933-1592.2007.00039x](https://doi.org/10.1111/j.1933-1592.2007.00039x).
- Kennett, Jeanette (2002) 'Autism, Empathy, and Moral Agency', *Philosophical Quarterly* **52**: 340–357.
- Kennett, Jeanette (2019) 'Competence, Attributability, and Blame: Resolving the Responsibility of the Psychopath', in Justin D. Coates and Neal A. Tognazzini, eds., *Oxford Studies of Agency and Responsibility (Vol. 5): Themes from the Philosophy of Gary Watson*: 142–164). Oxford University Press.
- Krizan, Zlatan and Anne D. Herlache (2017) 'The Narcissism Spectrum Model. A Synthetic View of Narcissistic Personality', *Personality and Social Psychology Review* **22**: 3–31. doi: [10.1177/1088868316685018](https://doi.org/10.1177/1088868316685018).
- Lamb, Matthew (2022) 'Quality of Will Accounts and Non-Culpably Developed Mental Disorders', *Journal of Ethics and Social Philosophy* **22**: 415–422. doi: [10.26556/jesp.v22i3.1421](https://doi.org/10.26556/jesp.v22i3.1421).
- Levy, Neil (2003) 'Cultural membership and Moral Responsibility', *Monist* **86**: 145–163.
- Levy, Neil (2009) 'Culpable Ignorance and Moral Responsibility. A Reply to FitzPatrick', *Ethics* **119**: 729–741.
- Miller, Joshua D., Hoffman, Brian J., Gaughan, Eric T., Gentile, Brittany, Maples, Jessica and Keith W. Campbell (2011) 'Grandiose and Vulnerable Narcissism. A Nomological Network Analysis', *Journal of Personality* **79**: 1013–1042. doi: [10.1111/j.1467-6494.2010.00711.x](https://doi.org/10.1111/j.1467-6494.2010.00711.x).
- Nelkin, Dana (2015) 'Psychopaths, Incurable Racists, and the Faces of Responsibility', *Ethics* **125**: 357–390. doi: [10.1086/678372](https://doi.org/10.1086/678372).
- Orth, Ulrich, Krauss, Samantha and Mitja D. Back (2024) 'Development of narcissism across the life span: A meta-analytic review of longitudinal studies', *Psychological Bulletin* **150**: 643–665. doi: [10.1037/bul0000436](https://doi.org/10.1037/bul0000436).
- Paulhus, Delroy L. and Kevin M. Williams (2002) 'The Dark Triad of Personality. Narcissism, Machiavellianism, and Psychopathy', *Journal of Research in Personality* **36**: 556–563. doi: [10.1016/S0092-6566\(02\)00505-6](https://doi.org/10.1016/S0092-6566(02)00505-6).
- Pereboom, Derk (2014) *Free Will, Agency, and Meaning in Life*. Oxford University Press.
- Pereboom, Derk (2021) *Wrongdoing and the Moral Emotions*. Oxford University Press.
- Pies, Ronald W. (2023) 'Narcissism, Empathy, and Moral Responsibility', *Philosophy, Psychiatry, and Psychology* **30**: 173–176. doi: [10.1353/ppp.2023.a899948](https://doi.org/10.1353/ppp.2023.a899948).
- Rolling Stone: Blame Game. Trump Says He Should 'Get All the Credit' if GOP Wins, 'Not Be Blamed' if They Lose. November, 8, 2022. <https://www.rollingstone.com/politics/politics-news/trump-credit-gop-wins-midterms-no-blame-lose-1234627288/>.

- Russell, Paul (1995) *Freedom and Moral Sentiment. Hume's Way of Naturalizing Responsibility*. Oxford University Press.
- Russell, Paul (2004) 'Responsibility and the Condition of Moral Sense', *Philosophical Topics* **32**: 287–305.
- Scanlon, Tim M. (1998) *What We Owe to Each Other*. Harvard University Press.
- Shoemaker, David (2017) *Responsibility from the Margins*. Oxford University Press.
- Shoemaker, David (2024) *The Architecture of Praise and Blame. An Interdisciplinary Investigation*. Oxford University Press.
- Smith, Angela (2008) 'Control, Responsibility, and Moral Assessment', *Philosophical Studies* **138**: 367–392. doi: [10.1007/s11098-006-9048-x](https://doi.org/10.1007/s11098-006-9048-x).
- Smith, Angela (2015) 'Responsibility for Attitudes: Activity and Passivity in Mental Life', *Ethics* **115**: 236–271. doi: [10.1086/426957](https://doi.org/10.1086/426957).
- Sripada, Chandra S. (2022) 'Impaired control in addiction involves cognitive distortions and unreliable self-control, not compulsive desires and overwhelmed self-control', *Behavioral Brain Research* **418**: 1–9. doi: [10.1016/j.bbr.2021.113639](https://doi.org/10.1016/j.bbr.2021.113639).
- Talbert, Matthew (2014) 'The Significance of Psychopathic Wrongdoing', in Thomas Schramme, ed., *Being Amoral: Psychopaths and Moral Incapacity*. 275–300. MIT Press.
- Talbert, Matthew (2016) *Moral Responsibility. An Introduction (Key Concepts in Philosophy)*. Polity Press.
- Talbert, Matthew (2022) 'Attributionist Theories of Moral Responsibility', in Dana Nelkin and Derk Pereboom, eds., *The Oxford Handbook of Moral Responsibility*. 53–70. Oxford University Press.
- Talbert, Matthew (2024) 'Moral Responsibility', in: Edward N. Zalta and Uri Nodelman, eds., *The Stanford Encyclopedia of Philosophy* (Fall 2024 Edition). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2022/entries/moral-responsibility/>.
- Urbonaviciute, Greta and Erica G. Hepper (2020) 'When is narcissism associated with low empathy? A meta-analytic review', *Journal of Research in Personality* **89** doi: [10.1016/j.jrp.2020.104036](https://doi.org/10.1016/j.jrp.2020.104036).
- Wallace, R. Jay (1996) *Responsibility and the moral sentiments*. Harvard University Press.
- Watson, Gary (1987/2004) 'Responsibility and the Limits of Evil. Variations on a Strawsonian Theme', in *Agency and Answerability. Selected Papers*: 219–259. Oxford University Press.
- Watson, Gary (1996/2004). 'Two Faces of Responsibility', in *Agency and Answerability. Selected Papers*: 260–288. Oxford University Press.
- Watson, Gary (2011). 'The Trouble with Psychopaths', in R. Jay Wallace, Rahul Kumar and Samuel Freeman, eds., *Reasons and Recognition: Papers on the Philosophy of T.M. Scanlon*: 307–331. Oxford University Press.
- Wolf, Susan (1990) *Freedom within reason*. Oxford University Press.
- Wolf, Susan (1987/2003) 'Sanity and the metaphysics of responsibility', in Gary Watson, ed. *Free will*: 372–387. Oxford University Press.