

# VIRTUE AND EXCUSE

Arden Ali

Northeastern University

October 4, 2018

(draft; please do not cite without permission)

This paper examines cases of wrongdoing that are excused because the agent acted under duress. Some philosophers explain these excuses by pointing to peculiar discrepancies in deontic norms. I argue that this approach is misguided. I sketch an alternative: agents who act wrongly under duress are excused because their wrongdoing is compatible with virtue in a sense to be described.

Some wrongdoings are excused. When someone offers an excuse for her wrongdoing, she admits that her act was wrong, but pleads that she is not fully blameworthy for performing it.<sup>1</sup> Cases of excused wrongdoing highlight a divergence in our judgments about the wrongfulness of acts and the blameworthiness of the agents who perform them.

Excuses in this sense are both familiar and diverse. Some excuses involve *ignorance*: a fiancée who mistakes cyanide for sugar may be excused for poisoning her partner's coffee, as long as her mistake is itself innocent. Other excuses involve brief *incapacity*: an employee who is provoked to anger, and temporarily loses his wits, may be excused for making a hateful remark about his antagonist. Even more kinds of excuse appeal to mere *accidents*: a father who endeavors to comfort his daughter may be excused if he inadvertently reminds her of a painful memory. The list could go on.<sup>2</sup>

---

<sup>1</sup> In this regard, my use of 'excuse' follows J. L. Austin; in cases of excuse, he says, "we admit that [the act] was bad but don't accept full, or even any, responsibility" (1956: 2). Also see Kelly (2013).

<sup>2</sup> Baron (2006) points to excuses grounded in civility (5) and excuses offered for future conduct (30), and others that are sometimes called exemptions (31). Tannenbaum (2014) discusses a class of partial excuses she calls 'mere moral failures.' Also see Husak (2005) and Duff (2006).

My focus in this essay is an especially puzzling class of excuses. These are not grounded in ignorance, incapacity, or mere accident. In cases of *duress*, clear-eyed agents decide what to do, and act wrongly, but they nonetheless appear at least partially blameless. To illustrate what I have in mind, consider the following two cases.

**Johann Vican** was living in Austria when the Nazis invaded in 1938.<sup>3</sup> Shortly after the Nazis arrived, Vican was expelled from Austria for ‘anti-Nazi activities,’ which included cutting down a German flag in protest of the occupation. While trying to return to Austria in 1940 to visit his mother, Vican was captured and imprisoned for three years. He was released on the condition that he join the SS and, although he agreed to the terms, he made no attempt to fulfill his side of the agreement after his release. He was consequently re-arrested and imprisoned at Dachau concentration camp, where he “was beaten several times a week and his weight dropped to about one hundred pounds” (Hilton 2004: 29). In 1944, he was moved to a sub-camp of Flossenbürg concentration camp, and suffered additional harms, including a poisoning that left him near death for weeks. The Nazis eventually made Vican the kapo of his group, which required that he participate in the administration and supervision of the camp. In this role, Vican obeyed orders to beat fellow inmates under threat that he would be punished or killed if he did not comply. Such orders were given at least fifty times. It is alleged that three unknown prisoners died at least partly because of Vican’s beatings in December 1944 and January 1945.

After the camps were liberated, Vican was taken into custody and charged with encouraging, aiding, abetting or participating in the killing of those three unknown

---

<sup>3</sup> See *United States v. Vican* (1947). The case is discussed in Heim (2013: 181–184)

prisoners. He pled guilty to these charges. During the trial, Vican's defense presented a statement from another kapo, who insisted that Vican acted only for the purpose of maintaining order and was not unnecessarily brutal (Heim 2013: 184). Commentators on the case have noted that Vican was not harmed or mistreated by the prisoners who were formerly under his supervision, even though they certainly had the opportunity to harm him and malevolent kapos were routinely attacked.

**Dražen Erdemović** was an ethnic Croat who lived in Bosnia.<sup>4</sup> He was a professed pacifist, but like all young men, he was required to complete his mandatory military service after his eighteenth birthday. His first stint in the military was short and unremarkable, even though the region had become engulfed in civil war by the time he left the service. After returning to civilian life, Erdemović tried desperately to find work outside the military, particularly as a locksmith, but he failed to secure reliable employment. He attempted to obtain visas to leave the region, but had little success. With his first child on the way, and no savings or forthcoming opportunities for work, Erdemović re-enlisted in the army. He requested that he be placed in a non-combat unit and was accordingly assigned to the 10th Sabotage Detachment, which focused mainly on maintaining specialized weaponry.

In July 1995, Erdemović's unit was dispatched to a farm. When they arrived, the soldiers were informed that they were expected to kill approximately one thousand Bosnian Muslims who had surrendered to the police and military. Erdemović immediately refused to take part in the murder and even argued with his superiors to spare an elderly man that he recognized in the group. Under threat that he would be

---

<sup>4</sup> See *Prosecutor v. Erdemović* (1998). Rosen (2014: 71–73) offers a more complete description of the case. Also see Yee (1997), Turns (1998) and Brooks (2003).

killed if he did not participate, Erdemović relented. He claims that he tried to kill as few people as possible and made an effort to spare the youngest victims. He later estimated that he directly killed about seventy people before being able to rally the support of three comrades and finally resist the order. In retribution for his insubordination, an unsuccessful attempt was made to kill Erdemović in the days following the incident.<sup>5</sup> After being treated for wounds sustained in the attack, Erdemović contacted a journalist and helped bring the atrocities of the war to international attention.

Some acts performed by Erdemović and Vican are deplorable. Vican beat his fellow prisoners and potentially caused the death of three of them; Erdemović killed nearly seventy people. But, when thinking about the blameworthiness of these acts, it seems unfair to disregard the fact that they acted under duress. Vican and Erdemović performed these terrible acts under threat of punishment or death. Although one might quibble about the details of these two actual cases, it seems to me that agents resembling Vican and Erdemović—people who knowingly and soberly perform wrong acts in terrible circumstances like theirs—are at least partially excused for their wrongdoing. By this, I mean that their blameworthiness seems mitigated by the circumstances in which they acted.<sup>6</sup> But this kind of excuse is quite puzzling. If ignorance, incapacity, or accident do not ground these excuses, what explains them?

My task in this essay is to offer an explanation of excuse of duress. In section one, I characterize the cases of interest more generally. Then, in sections two and three, I criticize two

---

<sup>5</sup> Yee (1997: 265)

<sup>6</sup> Although both crimes could have been punished with the death penalty, Vican and Erdemović were sentenced to five years and twenty years of imprisonment, respectively. See *United States v. Vican* (1947) and *Prosecutor v. Erdemović* (1998).

tempting but ultimately misguided approaches to the problem. In sections four, five, and six, I put forward and defend an answer: acts performed under duress are excusable because such actions are compatible with virtue in a sense to be discussed.

## 1. Duress as Excuse

Before I set out, I should say something about the crucial features of the kind of excuse I hope to explain.

First, in the relevant cases, the action performed is morally wrong.<sup>7</sup> Nearly everyone should agree that threats of punishment or death may sometimes *justify* behaviour that would otherwise be impermissible. Consider a bank teller who turns over his client's money under threat by an armed robber. Although the teller acts under duress, he doesn't act wrongly by handing over the money. His behaviour is justified by the fact of duress; when you are threatened by physical harm, it is typically permissible to give up valuable objects entrusted to you in order to protect your life. As I see it, things are different for Erdemović and Vican.<sup>8</sup> Unlike the teller, Erdemović and Vican act impermissibly in their respective circumstances. Given the gravity of their acts, and the sheer amount of suffering they produce, it seems Erdemović and Vican should not have acted as they did.

Of course, there is room to debate whether Vican and Erdemović really do act impermissibly. But this shouldn't concern us here. There must be some modification of their cases that makes their behaviour obviously impermissible. Claire Finkelstein suggests modifying cases like these by making the consequences of the acts even worse (2002: 331). At some point in this intensifying

---

<sup>7</sup> This is to say that Erdemović and Vican ought not to act as they do. The sense of 'ought' is both objective and subjective, since I assume that they have ample evidence about what their actions will bring about. I assume they violate an all things considered obligation, not a pro tanto one.

<sup>8</sup> Rosen discusses the idea the Erdemović's act may be justified in sections III and IV of his paper. See Rosen (2014: 73-78)

process, it will become clear that Vican and Erdemović should give up their lives instead of acting as they are directed. Still, Finkelstein thinks, these agents will be at least partially excused if they act under genuine threat of punishment or death. Likewise, in a study of Erdemović's case, Gideon Rosen suggests that we might modify the case such that "Erdemović could have saved some lives by standing with his victims and allowing himself to be shot" (Rosen 2014: 76). With this modified version of the case in mind, Rosen concludes: "*if there is ever a case in which a person might be morally required to die rather than kill the innocent, this is it, or this is close to it*" (2014: 78 his emphasis).

Second, in these cases of excuse, the agent acts deliberately. That is, the agents I have in mind *decide* to act as they do. It may have been the case that Vican and Erdemović were so paralyzed by fear that they briefly lost their capacity for reflective self-control.<sup>9</sup> But this needn't have been the case for them to deserve some kind of excuse. An agent who is excused by duress may have deliberated about what to do, at least as much as any of us do in everyday life. The ground for their excuse does not lie in the fact that they did not genuinely choose to act.<sup>10</sup>

Notice that, if agents like Vican and Erdemović deliberately perform a wrong act, they are not *compelled* in one familiar sense of the term. Consider Gary Watson's influential discussion of compulsion and weakness of will.<sup>11</sup> For Watson, weak-willed people "give in to desires which the possession of the normal degree of self-control would enable them to resist" but compulsive people give into desires "such that the normal capacities of resistance are or would be insufficient to enable the agent to resist" (1977: 48–49). He understands 'normal capacities of self-control' as those

---

<sup>9</sup> I borrow 'reflective self-control' from R. Jay Wallace. For Wallace, the powers of reflective self-control involve the ability to discern and apply moral reasons, as well as regulate one's behavior in the light of those reasons. See Wallace (2006: 86, 157)

<sup>10</sup> The chamber judges noted that "he was capable of taking positive action, once he had weighed up his options. The risks that he took appear to have been calculated and considered" (*Prosecutor v. Erdemović*, 1998). See Kahan & Nussbaum (1996: 333–338)

<sup>11</sup> Watson (1977)

capacities which are acquired in ordinary socialization and possessed by most adults (1977: 50–51). It is tempting to assimilate our cases of duress to cases of compulsion in Watson’s sense. However, if duress is expected to provide an authentic excuse for Vican and Erdemović, this temptation should be resisted. As Watson points out, weak-willed and compulsive agents are both constitutionally incapable of doing otherwise at the time of acting.<sup>12</sup> These agents differ because their incapacity is explained in different ways: the weak-willed person is incapable *because* she lacks normal self-control, but the compulsive person is incapable *in spite of* her normal capacity for self-control. But the person who acts wrongly under duress could have acted rightly. They are normally constitutionally capable of doing the right thing in their circumstances.

Third, in cases of excuse by duress, we blame agents less because they are less worthy of blame. This is to say that we do not merely pardon them. The difference between excusing and pardoning can be hard to see in any given case, but the two can be conceptually separated.<sup>13</sup> In order to pardon someone, one must judge that the person is blameworthy and yet choose to withhold blame. If I am late for a meeting for totally foreseeable reasons, I am fully blameworthy for my lateness, but my kindhearted colleagues may not blame me because they simply decide to let my tardiness pass. I take it that agents like Vican and Erdemović are excused rather than pardoned in this sense. We do not merely *choose* to free them from blame, but rather we judge that Vican and Erdemović *deserve* reduced blame.

Lastly, I assume that the blamelessness of Vican and Erdemović is not explained by our lack of *standing* to blame. Many conditions can apparently undercut our standing to blame another

---

<sup>12</sup> He says that a weak-willed person must be constitutionally able to meet standards of normal self-control, but may be “unable at the time of action to resist” (Watson 1977: 54). Similarly, a compelled person “is motivated by a desire (or ‘impulse’ or ‘inclination’) that he or she is unable to resist” (1977: 43). See Hieronymi (2007: 110).

<sup>13</sup> I use ‘pardon’ as Jeffrie Murphy uses ‘forgiveness.’ He writes: “We may forgive only what it is initially proper to resent; and, if a person has done nothing wrong or was not responsible for what he did, there is nothing to resent” (Murphy & Hampton 1988: 20). Also see Zaibert (2012).

person for a wrongdoing, including the fact that we would have done the same thing if we were in the person's circumstances.<sup>14</sup> But even someone with standing to blame should partially excuse those who act wrongly under duress. Imagine Erdemović's comrade, who resists taking part in the slaughter of prisoners, but miraculously survives subsequent execution attempts by his superiors. Since it is clear that this comrade would have given up his life to avoid any wrongdoing, he certainly has standing to blame Erdemović. Yet, it seems to me that even this comrade shouldn't fully blame Erdemović. Our task now is to explain why this is so.

## 2. Norm-splitting: Motivation

A simple Strawsonian explanation is tempting. Suppose that all blameworthy acts manifest *ill will*. Although Vican and Erdemović act impermissibly, they do not manifest ill will toward those they hurt, so Vican and Erdemović are (at least partially) excused for their actions. Despite its appeal, this proposal cannot really explain the cases in question. The supposition is false; it is plainly not true that all blameworthy acts manifest ill will.<sup>15</sup> Consider a blameworthy act of mere indifference. Suppose I promise to take care of my neighbor's cat over the weekend, but because I am preoccupied with one of my hobbies, I fail to do so. I am blameworthy for any harm that comes to the animal, it is not at all clear that I express ill will for the cat or its owner in behaving as I do.

---

<sup>14</sup> See Watson (1987: 239–254) and Wallace (2011: 326).

<sup>15</sup> Marcia Baron (2014) raises this point in the context of duress cases. She asks whether insufficient concern is a necessary condition on blameworthiness. If not, the fact that Vican and Erdemović have a sufficient level of concern does not explain their excuse. Any explanation of their excuse should point out a *necessary* condition on blameworthiness that they fail to meet. Baron's main counterexample is a case of *incompetence* where the agent does not manifest insufficient concern but still appears blameworthy (2014: 93). A patient dies during surgery because the tube providing her oxygen disconnects from the ventilator. Her anesthesiologist did not notice until an alarm signaled some kind of trouble. The doctor focused his attention on helping the patient, but because he hadn't noticed that the problem was a disconnected tube, he made all the wrong decisions. Baron submits that the anesthesiologist is blameworthy for his behavior, at least on some ways of spelling things out.



We might instead start with a different supposition: fully blameworthy acts manifest ill will or insufficient good will.<sup>16</sup> Though Vican and Erdemović act impermissibly, they may show sufficient good will toward those they hurt, so they are at least partially excused. Although the weaker supposition is likely true, this account must be further developed before it has any explanatory force. After all, Vican or Erdemović have an insufficient degree of good will in the most obvious sense—they don't care enough about others to act permissibly in their respective circumstances. To truly explain excuse by duress, one must say *why* their degree of good will is sufficient. It is certainly not enough to say that their degree of good will is sufficient to warrant excuse, since this is precisely what we intend to explain.<sup>17</sup>

Gideon Rosen suggests one way to develop this line of thought. He offers what I call a *norm-splitting solution*. In short, the proposal is that the deontic norms governing action and those governing moral concern come apart. As he puts it: “morality may require that one do A—e.g., sacrifice one's life—without requiring that one care enough about the underlying values to be *motivated* to do A” (2014: 84 his emphasis). Morality tells us, so to speak, that we are required to act

---

<sup>16</sup> See Strawson (1962: 76)

<sup>17</sup> At one point, Gideon Rosen suggests an explanation that seems to merely re-describe the phenomenon that we want to explain. He proposes that it is inappropriate to have attitudes of resentment toward Erdemović because, as he puts it, “Scanlon-style social distancing would not be warranted given the attitudes manifest in his act” (2014: 89). To socially distance oneself from another a human being in Scanlon's sense, one must withdraw some of the trust and willingness to cooperate that we ordinarily extend to others, even strangers. Blaming Erdemović is inappropriate because “he [is] entitled to a seat at the table, that is, to ordinary sociability” (Rosen 2014: 88).

The proposal is peculiar because distancing oneself from another person just is part of blaming that person in Scanlon's view. For him, “to blame a person is to judge [that the action shows something about the agent's attitudes toward others that impairs the relations that others can have with him or her] and to take your relationship with him or her to be modified in a way that this judgment of impaired relations holds to be appropriate” (2008: 128- 29). So, if one adopts Scanlon's view of blame, the proposal is circular. As a result, the proposed explanation cannot be neutral between analyses of blame. One might have hoped to explain excuse without taking a stand on what constitutes blame. In any case, as Rosen might agree, this explanation is not really a satisfactory stopping point. Resentment may be inappropriate because social distancing is inappropriate, but why is distancing inappropriate in response to wrongdoing performed under duress? Why is it inappropriate to revise one's relationship with Erdemović and Vican in response to what they did? It seems the proposal merely reframes the hard question about duress without answering it.

in a particular way but it is permissible to be unmoved to act in that way. If true, Rosen suggests, this schism would tell us why Vican and Erdemović have a sufficient degree of good will: they have a permissible level of moral concern.

Rosen admits that his norm-splitting proposal is merely a conjecture.<sup>18</sup> But it certainly matches some ordinary ways of thinking about duress and so it warrants careful attention. In what follows, I develop Rosen's conjecture a bit more fully and ultimately argue that it does not work.

Rosen's solution involves two deontic norms, one concerning action and the other concerning motivation. They might be described in the following way:

*Action Norm (AN)*                      E ought to do  $\phi$  in  $c$ .

*Motivation Norm (MN)*                E ought to be effectively-moved to do  $\phi$  in  $c$ .

The first deontic norm, AN, is the familiar norm governing what we do. The dictates of this norm often take the form of imperatives like 'Take the trash out when you have promised to do so.' But we might instead put these dictates in the form of declarations about a particular agent's obligations, e.g. 'Stephen ought to take the trash out when he promised to do so' or, more generally, 'E ought to do  $\phi$  in  $c$ .'

The second deontic norm, MN, is less familiar. This is partly because it includes unfamiliar terminology, i.e. 'effective motivation,' and partly because it is somewhat strange to say that someone is obliged to be motivated. Rosen sees the latter concern, but he doesn't find it troubling. He says: "It makes perfectly good sense to ask whether gratitude, for example, is morally required ('owed') in a given case. The fact that we are passive with respect to our sentiments may entail that when we have done what we can by active means to comply with our obligations in the realm of feeling, we are

---

<sup>18</sup> Rosen (2014: 85–86)

then blameless for our failures; but that is consistent with our having genuine moral obligations in this realm" (Rosen 2014: 84-85). For the sake of argument, suppose that Rosen is correct on this point; it is intelligible to talk about our obligations to be motivated, i.e. our obligations to desire or feel.

The unfamiliar terminology is another matter. MN refers to ‘effective motivation’ rather than motivation simpliciter. As I use the term, someone is effectively-moved to perform an act if and only if she is motivated in such a way that she in fact tries to perform the act. I gather that MN couldn’t be a principle about how one ought to be motivated simpliciter. To see why, we first need to put Rosen’s full explanation in view. His explanation has two steps:

*Rosen’s norm-splitting solution:*

Step 1: For some cases  $c$ , E is required to do  $\phi$ , but E is not required to be effectively-moved to do  $\phi$ .

Step 2: In  $c$ -cases, if E is motivated to do  $\phi$  to degree  $d$  but fails to do  $\phi$ , E is at least partially excused.

The first step is just a restatement of Rosen’s claim that “morality may require that one do A—e.g., sacrifice one’s life—without requiring that one care enough about the underlying values to be *motivated* to do A” (Rosen 2014: 84). The second step states that agents are excused for wrongdoing when this split occurs, as long as they are motivated to act permissibly to some extent. It adds that the agent must be motivated to act permissibly to some extent (i.e. ‘to degree  $d$ ’) in order to avoid becoming too liberal. Without this qualification, *anyone* who finds themselves in Erdemović’s situation will be excused. Erdemović’s sadistic comrade may know that his commander

will kill him if he does not comply. But, if this sadistic comrade has absolutely no qualms about murdering civilians and so kills the civilians happily, he should not be excused.<sup>19</sup> Since the value for *d* is probably nontrivial (or so Rosen would insist), such a person would not be excused on the view described above.

Now it may be clear why MN refers to effective motivation instead of motivation simpliciter. If MN were just about motivation simpliciter, there must be cases of required acts that agents are not required to be moved to do *at all*. I suspect that this is a stronger claim than the one Rosen wants to defend. Intuitively, it is impermissible for Erdemović to be totally unmoved to spare the civilians; it is also impermissible for Vican to be totally unmoved to spare his fellow prisoners. Rosen relies on the more plausible claim that Erdemović and Vican are required to be motivated to do the right thing *to some extent*, but they aren't required to be motivated in such a way that they in fact try to perform it. That is, they are not required to be effectively-moved.

Rosen's version of norm-splitting fails at Step 1. That is, AN and MN do not come apart in the way he proposes; if you are required to perform an act, you are also required to be effectively-moved to do it. Notice that it is difficult to provide any examples that clearly show these norms coming apart, putting aside the cases Rosen intends to explain. Admittedly, it is sometimes permissible to be totally unmoved to act even though it would be morally *good* to act. Typical examples are acts of heroism or generosity. Saving a stranger from a burning building at great personal risk is morally good, but plausibly it is permissible to be totally unmoved to do it. Giving a significant portion of one's income is also morally good, but plausibly it is permissible to be totally unmoved to do that too. These examples don't help the norm-splitting view under consideration. Whenever it is permissible to be totally unmoved to act—as in heroism or generosity—the act is not morally required. This is tantamount to the claim that the cases above are cases of *supererogation*.<sup>20</sup>

---

<sup>19</sup> Compare the third case in Frankfurt (1969: 832).

<sup>20</sup> See Urmson (1958) for the seminal discussion of supererogation.

It is true by definition that agents are not obligated to perform supererogatory acts. For this reason, it is a mistake to treat cases of excuse as failures to perform supererogatory acts. In cases of excuse, there must be wrongdoing, but in cases of supererogation there is no wrongdoing.

More substantively, I think the schism between AN and MN is incompatible with two general principles about normative reasons. In short, a normative reason is a consideration that is capable of moving *and* obligating an agent to act. In that way, the concept of a reason connects obligation and motivation inextricably. If this is true, AN and MN cannot come apart, so Rosen's proposal fails. To see how this argument works in detail, consider the following two principles:

*Justification:* If I am obligated to do  $\phi$  and not obligated to be effectively-moved to do  $\phi$  in the same circumstances, there must be some good reason to do  $\phi$  that is not also a good reason to be effectively-moved to do  $\phi$ . Call this reason  $p$  (or the fact that  $p$ ).

*Internalism:* If the fact that  $p$  is a good reason for E to do  $\phi$ , then E is capable of being moved to do  $\phi$  by the belief that  $p$ .

The first claim, called *Justification*, is a principle about the connection between obligations and normative reasons. It says roughly that obligations are grounded in normative reasons. That is, if I am obligated to perform some action or have some motivational state, there are good reasons in virtue of which I have these obligations. Accordingly, if E is obligated to do  $\phi$  in  $c$ , but not obligated to be effectively-moved to do  $\phi$  in  $c$ , there must be some good reason in virtue of which E has the former obligation but not the latter.<sup>21</sup>

---

<sup>21</sup> Couldn't there be a normative reason *against* being moved to do  $\phi$  that is not a normative reason against doing  $\phi$ ? I did not spell things out in this way to avoid confusion. But the upcoming argument can be recast in this form. If the fact that  $p$  is a normative reason against being moved to do  $\phi$ , then the fact that

The second claim, called *Internalism*, concerns the connection between reasons and motivation. It says roughly that normative reasons are necessarily capable of motivating action. If the belief that  $p$  couldn't in any way move E to do  $\phi$ , then the fact that  $p$  is simply not a reason for E to do  $\phi$ .<sup>22</sup> Although this principle is the subject of substantive debate in the philosophy of action, it is widely accepted.<sup>23</sup>

In conjunction with *Justification* and *Internalism*, Rosen's proposal leads to an absurd result. Here's how. The following claim follows from Rosen's view and *Justification*:

- (1) The fact that  $p$  may be good reason for E to do  $\phi$  even though  $p$  is not a good reason for E to be effectively-moved to do  $\phi$ .

Rosen says that Erdemović is obligated to resist the order but not obligated to be effectively-moved to resist it. So, granting *Justification*, Erdemović must have some good reason to resist (namely the fact that  $p$ ) that is not a good reason to be effectively-moved to resist.

But now consider the perfectly rational agent in Erdemović's situation. The perfectly rational agent is constituted such that she is responsive only to good reasons. In other words, if the fact that  $q$  is a good reason to do  $\psi$ , the perfectly rational person is constitutionally capable of being moved to do  $\psi$  by the belief that  $q$ ; if the fact that  $q$  is not a good reason to do  $\psi$ , then the perfectly rational

---

$p$  is a normative reason in favour of being moved to do not- $\phi$ . The perfectly rational person will remain in an impossible position. On one hand, she must be capable of being moved to do not- $\phi$  by the belief that  $p$ . This follows from *Internalism* and the assumption that the fact that  $p$  is a normative reason in favour of being moved to do not- $\phi$ . But, on the other hand, the perfectly rational person is incapable of being moved by bad or irrelevant reasons for action. So, if the fact that  $p$  is also not a normative reason to do not- $\phi$  (either a bad reason or an irrelevant one), she is incapable of being moved to do not- $\phi$  by the belief that  $p$ .

<sup>22</sup> Most internalists read the 'could' in this principle quite liberally. For Bernard Williams, it will still be true that someone has a reason even if they could only be moved after robust exercises of imagination and reflection. See Williams (1981a: 101-114).

<sup>23</sup> For discussion of internalism and its motivations, as well as some reasons for doubt, see Setiya (2012: 1-14)

person is constitutionally incapable of being moved to do  $\psi$  by the belief that  $q$ . I take it that this psychological constitution is possible.

According to (1), the fact that  $p$  is not a good reason to be effectively-moved to do  $\phi$ . If the fact that  $p$  is not a good reason to be effectively-moved to do  $\phi$ , then the perfectly rational agent is constitutionally incapable of being moved to be effectively-moved to do  $\phi$  by the belief that  $p$ . This claim can be simplified: the perfectly rational agent is constitutionally incapable of being moved to  $\phi$  by the belief that  $p$ .<sup>24</sup>

But, according to (1), the fact that  $p$  is *also* a good reason to do  $\phi$ . From *Internalism* and (1), it follows that the perfectly rational agent is capable of being moved to do  $\phi$  by the belief that  $p$ . This is because the fact that  $p$  is a good reason for her to do  $\phi$  and normative reasons are necessarily capable of motivating the person whose reasons they are. The result is two incompatible claims:

(2) The perfectly rational person is constitutionally incapable of being moved to do  $\phi$  by the belief that  $p$ .

(3) The perfectly rational person is constitutionally capable of being moved to do  $\phi$  by the belief that  $p$ .

The conjunction of (2) and (3) is an absurdity. It means that there is no coherent way for the

---

<sup>24</sup> Is this simplification fair? I think so. For those who are skeptical, notice that the argument against Rosen's proposal can be run without making it. If the fact that  $p$  is not a good reason to be effectively-moved to do  $\phi$ , there is no deliberative route from the perfectly rational person's existing attitudes (and the belief that  $p$ ) to motivations required to perform the act. Given her constitution, it is impossible to devise a 'psychological link' from her attitudes and the belief that  $p$  to her performance of doing  $\phi$ . The fact that  $p$  therefore couldn't be a motivating reason for the perfectly rational person; that is, she couldn't do  $\phi$  on the ground that  $p$ . If she is perfectly responsive to good reasons, then obviously the fact that  $p$  cannot be a good reason to do  $\phi$ . This conflicts with (1), according to which the fact that  $p$  is a good reason to do  $\phi$ .

perfectly rational person to be constituted. She is supposedly both susceptible and insusceptible to doing  $\phi$  by the belief that  $p$ . To avoid this result, one could reject *Justification* or *Internalism* or the idea of a perfectly rational agent. But I gather that these items are far more plausible than Rosen's purported split between AN and MN. Since Rosen's solution to duress cases depends on this split, the proposal fails.

Rosen implies that someone like Erdemović doesn't have sufficient reasons to be moved to act even though he has sufficient reasons to act. I claimed that normative reasons must be capable of motivating the agents whose reasons they are, so sufficient reasons to act *just are* sufficient reasons to be moved to act. The commonsense view is correct: if one is obligated to act, one is also obligated to be effectively-moved to act.

There is an important objection facing anyone who denies Rosen's style of norm-splitting. To see the objection, imagine a person who suffers from severe claustrophobia. She experiences uncontrollable fear, anxiety, and panic whenever she is enclosed in small spaces. Suppose this claustrophobic person sees an unsupervised child run into an alley and crawl into a ventilation duct. Given the danger for the child, the claustrophobic person may be required to intervene even though it is *very difficult* for her to motivate herself to crawl into the ventilation duct. Compare someone with no such disorder in the same circumstances. A non-claustrophobic person is required to intervene to protect the child too, but unlike the claustrophobic person, she will find it *very easy* to motivate herself to crawl into the space.

This contrast generates the objection. I argued earlier that AN and MN do not come apart. If I am right, and these two people are required to intervene on behalf of the child, they are also required to care enough about the child that they in fact try to act. But then I am committed to saying that the claustrophobic person is required to care about the unknown child *more* than the non-claustrophobic person is. After all, it is hard for the claustrophobic person to act, but easy for her counterpart. Presumably, if acting is harder for someone with claustrophobia, greater concern about



the child is needed in order to act. There is something apparently unfair about this. One might think morality should require that people care about strangers to exactly the same extent.

The preceding line of argument brings out what is most appealing about Rosen's proposal. But ultimately I think it is mistaken. The objection points out that my position seems to generate an inequality in the amount of moral concern required from different people. But the mere fact that two people are required to have different levels of moral concern is not always strange or objectionable. Suppose I am walking around East Lake, and you are walking around West Lake, and by pure coincidence, there is a person drowning in both. If I am wearing my favorite outfit, and you do not care at all about your clothing, it will be slightly harder for me to jump into East Lake than it is for you to jump into West Lake. It will require slightly more moral concern for me to intervene than it will for you. But morality dictates that we should both be moved to act even though acting requires more concern from me than it does from you. I take it that there is nothing bizarre or objectionable about this inequality. This is because the inequality in the East/West Lake case is generated by something relatively unimportant. That is, when people's lives are at stake, my love for an outfit is too trivial to matter. In general, it is acceptable to require unequal amounts of moral concern from different agents, provided that the reasons that generate the inequality do not outweigh the reasons to act.

Now the objection can be dismissed. If the principle above is right, then inequality in the required degree of concern is *always* acceptable when agents are all-things-considered required to act. Why? To say that someone is all-things-considered required to do  $\phi$  is to say that they ought to do  $\phi$ , given *all* the facts. This means that any fact that might generate unequal levels of required concern must be trivial or otherwise unable to outweigh the reasons to do  $\phi$ . If there were some fact capable of outweighing the reasons to do  $\phi$ , then presumably that fact would have blocked the all-things-considered obligation to do  $\phi$  in the first place. As a result, if an act is all-things-considered required, it simply does not matter that two people may be required to have different levels of

concern.

Take the claustrophobia example. To the extent that one insists that the claustrophobic person is required to act, one also must insist that it is acceptable to require greater moral concern from her in this case. In other words, if her claustrophobia is so extreme that she needs an *unbelievable* amount of moral concern to act, then her claustrophobia is usually so extreme that it releases her from her obligation to act. And my opponent must insist that the claustrophobic person is required to act because otherwise there is no parallel to the case of excused wrongdoing. Thus the worry raised by the claustrophobia example can be resolved without splitting the deontic norms governing action and motivation in the radical way that Rosen proposes.

### 3. Norm-splitting: Character

Hume's remarks about blameworthiness inspire a different version of norm-splitting.<sup>25</sup> Hume claimed that wrong acts are not blameworthy unless they are rooted in morally poor dispositions. Several philosophers have endorsed this principle, occasionally touting its ability to deal with excuses like duress.<sup>26</sup> As usually interpreted, Hume offers a norm-splitting explanation of excuse. But his version of norm-splitting doesn't involve the deontic norm governing motivation (i.e. MN). Instead, it concerns the norms governing permissible *general dispositions*. Consider the following two norms:

*Action Norm (AN)*                      E ought to do  $\phi$  in *c*.

*Character Norm (CN)*                      E ought to be disposed to do  $\phi$  in all *c*-like cases.

---

<sup>25</sup> See his *A Treatise of Human Nature*, book II, part III, sec. II.

<sup>26</sup> See, for example Brandt (1958: 3–29), Hobart (1966), Bayles (1982), Vuoso (1987), and Finkelstein (2002). For discussion, see Sher (2007: 18–32).

AN is familiar. But the character norm (CN) requires some unpacking. It is natural to put the dictates of CN in terms of virtues, e.g. ‘Stephen ought to be grateful.’ If we understand the virtues as dispositions, this norm can be restated as ‘Stephen ought to be disposed to thank others when they have helped or supported him’ or, generally, ‘E ought to be disposed to do  $\phi$  in  $c$ .’ Since AN specifies an act that someone must perform in a particular case, the corresponding character norm specifies that the person must be disposed to act in all cases that are similar, morally-speaking, to the action required by AN.

Of course, it is difficult to say which cases are morally-similar, or  $c$ -like. But if one takes the idea that there are virtues and vices seriously, one is committed to the existence of *some* similarity relation between cases. All cases that call for honesty bear this relation to one another. All cases that call for courage bear this relation to one another, and so on. The individuation of the virtues is messy, but I assume the mess can be sorted.<sup>27</sup> At the very least, such an assumption is needed to get the Humean proposal off the ground.

How does CN differ from MN? Recall that MN is a norm about how one ought to be effectively-moved to act in a particular case. CN is a norm about how one ought to be disposed to act in all circumstances of that type. An agent might satisfy MN without satisfying CN. Imagine a corrupt politician who is only inclined to tell the truth when honesty serves him in some way. One day, he tells the truth about an ongoing scandal because lobbyists have offered him a large sum of money to do so. The politician does the right thing and thereby satisfies AN. He is also effectively-moved to do the right thing and so satisfies MN. But he isn’t disposed to tell the truth in most cases where truth-telling is required. So he fails to meet CN.

Conversely, an agent might satisfy CN without satisfying MN. This occurs when good people

---

<sup>27</sup> For one sorting scheme, see Russell (2009: 177-209). For an empirically-informed discussion of the individuation of the virtues, see Doris (2002).

are ignorant about some matter of fact. The perfectly considerate person is disposed to take care in her treatment of others in *every* circumstance. Suppose this person lacks some information—say, she doesn't know that there is a colleague walking behind her—and thus acts wrongly by slamming the door in her colleague's face. The considerate person is disposed to hold the door open, but since she is oblivious to her colleague's presence, she slams the door. In a case like this, the person is disposed to act well in all cases where considerateness is called for, but she isn't effectively-moved to do the right thing. She meets CN but fails to meet MN.

In short, the difference between MN and CN is that MN tells us how much one ought to be motivated to do the right thing, but CN tell us how one must be generally disposed in order to count as virtuous in some respect. MN concerns morality's demands on our motivation in particular cases, but CN concerns morality's demands on our general character.

In Rosen's proposal, AN and MN come apart. In Hume's proposal, AN and CN come apart. This means that the virtuous person needn't be disposed to fulfill *all* her obligations. Someone can still count as honest even though she is disposed to lie in some cases where honesty is required. Someone can still count as brave even though they are disposed to act cowardly in some cases where bravery is required, and so on. This claim is the starting point for Humean norm-splitting. In full, the proposal is:

*Humean norm-splitting solution:*

Step 1: For some case *c*, E is required to do  $\phi$  in *c*, but E is not required to be disposed to do  $\phi$  in all *c*-like cases.

Step 2: If E is disposed to do  $\phi$  in typical *c*-like cases, but fails to do  $\phi$  in an atypical *c*-like case, E is at least partially excused for doing  $\phi$  in the atypical *c*-like case.

Step 1 is the idea that AN and CN come apart. Step 2 gives the sufficient condition for excuse. In cases where AN and CN come apart, an agent is excused for her wrong act, provided that she is disposed to act well in typical *c*-like cases. To explain Vican's case, one might claim that Vican is obligated to resist injuring his fellow prisoners, but nonmaleficence doesn't require that he is disposed to resist in his terrible circumstances. As long as Vican is disposed to act nonmaleficently in typical cases, he deserves an excuse for failing to resist the guard's orders. Likewise for Erdemović. He is obligated to resist his superior's orders, but nonmaleficence doesn't require that he is disposed to resist in his atypical circumstances. If Erdemović is disinclined to harm others in typical cases, he too deserves an excuse for his wrongdoing.

Does Humean norm-splitting work? The explanation obviously depends on the assumption that Vican and Erdemović are disposed to act permissibly in the relevant typical cases. I think this assumption is plausible given their behaviour in other situations. It is hard to dismiss the claim made in Step 1 of the proposal on purely formal (as opposed to substantive) grounds. It might well be that our standards for action and character diverge. People might be benevolent, courageous, or honest, even though they are not disposed to act well in *all* of the relevant cases. I see no theoretical hurdle standing in the way of such a view.

The sufficiency claim, Step 2, in the Humean norm-splitting solution is more suspicious. It faces a familiar objection. Why should the mere fact that the agent *would* act well in other circumstances mitigate his blame for manifesting a poor disposition in his *actual* circumstances?

George Sher develops this objection to the sufficiency claim.<sup>28</sup> He describes an agent named Alphonse, who is "a generous and empathetic person" and who "tries scrupulously to do the right

---

<sup>28</sup> George Sher (2007: 29) discusses this objection to character theories of excuse in general. Also see Moore (1997: 579) and Smith (2008).

thing” (2007: 26). But Alphonse has a moral blind spot: he judges himself extremely harshly for his rare lapses in moral perfection, and when judging himself, he tends to behave “in a spirit of moral self-defilement” (2007: 24). Sher goes on to describe a particular episode in Alphonse’s life: on one occasion, Alphonse believes that he has made a moral mistake and begins to morally self-destruct by making “a cruel remark of precisely the kind that he most abhors” (2006: 23–24). Although Alphonse is disposed to act well in nearly every circumstance, he is disposed to act poorly on the exceptionally rare occasion that he makes a moral mistake.<sup>29</sup> Sher thinks this case is a counterexample to the sufficiency claim made in Hume’s norm-splitting proposal. By the lights of the proposed explanation, he deserves at least partial excuse for his wrongdoing. After all, Alphonse is disposed to act well in all typical cases, even though he is disposed to act cruelly in his current circumstance (and does in fact act cruelly). But Sher’s intuition is that Alphonse is fully blameworthy for making his cruel remark. People will often be blameworthy for acting badly in circumstances where they are disposed to act badly, regardless of how rare these cases happen to be.

The details of Alphonse’s case can be ignored. Sher’s point is that counterexamples to the Humean proposal do not require any stretch of the imagination. The formula for a counterexample is to describe someone who is disposed to act well in typical circumstances, but disposed to act viciously in some (perhaps rare) cases. When those atypical cases come to pass, the person is blameworthy for what they do, but Hume is committed to excusing them. A teenager might act kindly in most typical circumstances, but turn extraordinarily cruel when stressed by her college applications. A pastor might be compassionate and empathetic to nearly everyone he meets, but become unjustifiably ruthless when confronting the person who justifiably hurt his child. In cases like these, the agent’s blameworthiness does not seem blocked simply by the person’s disposition to

---

<sup>29</sup> Sher notes that “the dispositions that come together to produce Alphonse’s cruel remark...are not themselves dispositions to act cruelly; and neither (we may suppose) would their interaction lead Alphonse to act cruelly under many other circumstances” (2007: 24–25).

behave differently in typical circumstances.<sup>30</sup>

#### 4. Fallible Virtue

Norm-splitting is one approach to the puzzle posed by cases of duress. A norm-splitter identifies some discrepancy in deontic norms and then uses this discrepancy to explain cases of duress. I have argued that two ways of developing the approach fail. I am doubtful that any norm-splitting proposal will work.

Where does this leave us? It leaves us in need of an alternative to norm-splitting. One candidate explanation focuses on the fact that both Vican and Erdemović experience psychological resistance. They *feel* pressured when carrying out the orders of their superiors. The pressure may not be so unbearable that it disrupts self-control to such an extent that they no longer act deliberately. But the pressure may nonetheless be substantial. If so, their excuse may seem to lie in the phenomenology of acting under duress: Vican and Erdemović are excused because they experience substantial psychological resistance as they act.

As one might expect, this simple psychological explanation faces fatal objections. First, psychological resistance may have immoral sources. Suppose a soldier feels resistance to carrying out horrible orders, but only because he really desires to carry out some more horrible acts instead. Such a person would have the experience of resistance, but their resistance is rooted in something abhorrent, and so they shouldn't be excused. Second, if the experience of resistance is sufficient for an excuse, then those who feel psychological pressure when the threat is not genuine or not "important on some scale of values independent of the agent's own" deserve excuses as well

---

<sup>30</sup> Here I emphasize 'simply' because I agree that a person's broader disposition is relevant to her blameworthiness for a particular act. But the *mere* fact that she is disposed to act well in other cases doesn't seem sufficient for an excuse when she is disposed to behave badly in her current circumstances.

(Richards 1987: 23). Suppose an agent commits murder under threat that the last remaining slice of cake would be eaten if he did not. This threat might cause him distress if he really cares about the final slice of cake, so he might commit the murder while experiencing the same kind of resistance that unsettles Vican and Erdemović. But clearly the threat is too trivial to generate an excuse.

In the light of these objections, norm-splitting looks appealing once again. With psychological views in mind, Jeffrie Murphy concludes that, “in most cases properly called duress...what we are characterizing is...something mainly moral rather than mainly psychological” (1992: 133). But I think the move to norm-splitting is too quick. In one way, psychological views are on the right track: the experience of intense inner struggle is crucial when we excuse agents because they acted under duress. If, for example, Erdemović did not experience considerable inner struggle while carrying out the murder, I gather our judgment that he is excused would subside. This pushes in the direction of an approach according to which the explanation of duress is neither purely psychological nor purely normative. Before I can set out this alternative, I first need to set up a distinction between what I call *full* and *infallible* virtue.

According to a familiar picture, virtue varies along a single dimension. Agents who fall on one end of the spectrum are completely virtuous and agents who fall on the other are completely vicious.<sup>31</sup> Barring ignorance, those who are completely virtuous always do what morality asks and those who are completely vicious never comply. I propose a shift away from this picture of virtue. I claim that there are at least two dimensions along which virtue varies, which I call full virtue and infallible virtue.

Someone is *fully virtuous* when she possesses the disposition to act permissibly. More helpfully, it is a matter of having the beliefs, desires, hopes, aversions, tastes, and knowledge that ground a disposition to act permissibly. There is room for a view that is psychologically rich,

---

<sup>31</sup> See, for example, Annas (2004)



ascribing many complex cognitive and emotional states to the virtuous person, and there is room for a view that is psychologically meager, ascribing only want and aversion to her.<sup>32</sup> Virtuousness in this sense may come in degrees; a fully virtuous person is disposed in exactly the way moral theory prescribes, but a less fully virtuous person matches this disposition incompletely.

Someone is *infallibly virtuous* when she always acts permissibly (unless she is mistaken about some matter of fact). If full and infallible virtue come apart, then it is possible to always act permissibly, and so be infallibly virtuous, but lack a disposition to act well, and so fail to be fully virtuous in my sense. I take it that this possibility is real but quite unlikely.<sup>33</sup> If a person reliably does the right thing, it is usually the case that she has the disposition to act permissibly. But the connection could be severed. Imagine an evil person who always tries to do the wrong thing, but unwittingly always ends up doing the right thing. This person may always act permissibly, but (we may assume) he lacks any of the beliefs, desires, and knowledge that usually characterize good people. The incompetently-evil person is infallibly but not fully virtuous.

The distinction between full and infallible virtue also implies that it is possible to maintain a virtuous disposition, but fail to act permissibly on an occasion. This possibility appears to face a more serious conceptual problem. In the remainder of this section, my goal is to show that this problem can be resolved.

When someone is fully but fallibly virtuous, she is disposed to act permissibly but sometimes fails to act permissibly. Our ordinary practice of ascribing dispositions allows for this possibility. In general, dispositions survive some failures to manifest. A professional basketball player may be disposed to make a free throw when she tries. But, she may miss in spite of her disposition to make

---

<sup>32</sup> Julia Annas (2011) provides a psychologically rich picture of virtue. Richard Brandt (1970) gives a notably simpler account.

<sup>33</sup> The claim here is actually stronger than what is needed. As I defined the terms, full virtue is neither necessary nor sufficient for infallible virtue. The ultimate explanation of duress requires only that full virtue is not sufficient for infallible virtue, i.e. someone can be disposed to act permissibly, but fail to act permissibly.

the shot. In these cases, one might say that the disposition *tolerates* the failure.

The idea that psychological dispositions may tolerate failures to manifest has been discussed elsewhere.<sup>34</sup> Notably, John Heil and C. B. Martin appeal to fallibility in order to circumvent one of Kripke's objections to the dispositional analysis of rule-following.<sup>35</sup> Kripke claimed that grasping a rule cannot be a matter of having a disposition because of the possibility of mistake. According to the objection, the dispositional account cannot distinguish someone who makes a mistake in applying a rule and someone who grasps a different rule altogether. In response, Heil and Martin claim that an agent's "having acquired, and continuing to possess, the disposition does not mean that [he] will manifest it on every occasion in which its manifestation is called for" (1998: 291). In cases of mere mistake, the agent's rule-following disposition is blocked or inhibited but nonetheless still present.

Likewise, D. M. Armstrong appeals to fallibility to defend his claim that having an experience that *p* requires being disposed to believe that *p*. In cases of illusion, the world appears to an agent in a peculiar way, but she may withhold believing that things are how they appear. In short, Armstrong's reply is that the appearance still "pushes towards being a belief, but is held back by other contradictory beliefs" (1961: 106).

The main objection facing the idea of full but fallible virtue proceeds as follows. If full but fallible virtue is possible, then someone who is disposed to act permissibly may act impermissibly on some occasion. On one hand, if a given failure act permissibly is explained by a feature external to the agent—say, she is physically prevented of fulfilling her intentions—then the agent does not *truly* act impermissibly. After all, if she is prevented from performing an act, she cannot be obligated to perform it. But, on the other hand, if the disposition to act permissibly is interrupted by some feature internal to her psychology—say, her beliefs or desires—then she must lack the psychological

---

<sup>34</sup> There is evidence that Thomas Aquinas endorsed the distinct possibility of fallible dispositions to act. On this point, see Kent (2012: 151–154).

<sup>35</sup> Heil & Martin (1998: 291) and Kripke (1982). See also Bird & Handfield (2008).

disposition in the first place.<sup>36</sup> So, it seems, one cannot have a disposition to act permissibly and act impermissibly.

It seems to me that an agent can fail to manifest one of her dispositions on account of internal features of her psychology. Consider a case where an agent's disposition to act is interrupted by the agent's mental behaviour. Emerson is an extraordinarily anxious person. At any mention of a performance review at his job, he is disposed to engage in a series of pathological behaviours, e.g. recalling his many mistakes, wondering intensely about the review, and deliberating about what to do if he is fired. Knowing that he is disposed in this way, he decides to rehearse a favorite song to distract himself whenever the review is mentioned. Emerson remains anxious and disposed to perform unhealthy mental acts despite his failure to do so when he rehearses the song. His anxious disposition makes his mental rehearsing intelligible; he sings to himself because he is prone to anxious spirals. It seems, intuitively, that Emerson has interrupted his disposition by performing a mental act. His failure to act is explained entirely by a feature internal to his psychology.

I should note that Emerson's case may be an example of *intrinsic masking*. Although few doubt that dispositions can have extrinsic masks, there is considerable disagreement about the possibility of intrinsic masks in metaphysics.<sup>37</sup> Consider a standard case of extrinsic masking: a fragile glass wrapped in bubble wrap may not break when struck.<sup>38</sup> The wrapped glass remains disposed to break when struck; its disposition remains but its manifestation is masked by the wrapping material. Some philosophers believe it is crucial that the masking property—being covered in bubble wrap—is an extrinsic property of the glass. To them, it would be absurd to say that an

---

<sup>36</sup> Bird & Handfield (2008) raise a version of this objection to Heil & Martin's proposal.

<sup>37</sup> In addition to cases of masking, there are also cases of *finking*. Take David Lewis' well-known example of finking: a sorcerer protects a fragile glass by carefully watching over it and quickly casting a spell that makes the glass unbreakable if it is struck. In this case, the glass' disposition to break when struck is finked by the sorcerer. For clarity, I focus mostly on masking in the main text, but cases like Emerson's could be understood in terms of finks as well.

<sup>38</sup> See Johnston (1992) and Bird (1998).

intrinsic property can mask the manifestation of disposition in the way that the bubble wrap masks the glass' disposition to break. Suppose the sorcerer decided to “enchant the glass in such a way that it has an intrinsic property which would cause it, when struck, to lose the molecular structure which, under normal circumstances, is the casual basis of fragility” (Bird & Handfield 2008: 8). One might insist that a glass with this peculiar intrinsic property is simply not a fragile glass.

Settling the debate about intrinsic masking is beyond the scope of this paper.<sup>39</sup> There may be space for a hybrid view, according to which intrinsic masking is possible only when the purported masker is temporary or fleeting, just as caffeine can temporarily block the manifestation of one's disposition to fall asleep when tired.<sup>40</sup> Even further, it may be possible for an attitude or cognitive process to be internal to someone's psychology without being intrinsic to the agent, in which case intrinsic masking needn't be invoked at all.<sup>41</sup> For my purposes, it is enough to say that Emerson's case demands some kind of explanation even from those who deny the existence of intrinsic masking. The point needn't rest on any particular analysis or explanation of Emerson's case. If the case is possible, it is also possible to be fully but fallibly virtuous in exactly the same way.

In this section, I have introduced and briefly defended a distinction between full virtue and infallible virtue. Many philosophers note that virtue is not the exclusive property of virtue ethics.<sup>42</sup> Even crude utilitarianism (for example) offers a conception of virtue; for a very simple utilitarian, the virtuous person has *whatever* psychological composition is required to ensure that she maximizes overall utility. To be fully virtuous, a utilitarian claims, one must possess the disposition that does not tolerate any failures to maximize utility. But a virtue theorist might say the utilitarian's account of

---

<sup>39</sup> For skepticism about intrinsic finks and masks, see Handfield (2008: 302), Handfield & Bird (2008), and Choi (2009; 2012). For defenses of intrinsic finks and masks, see Clarke (2008), Fara (2008), and Ashwell (2010, 2015).

<sup>40</sup> See Fara (2008: 468) for this example.

<sup>41</sup> Some argue that mental states and cognitive processes can be embedded outside the biological brain. These attitudes or processes may be part of someone's psychology without being intrinsic to the agent. See Clark & Chalmers (1998).

<sup>42</sup> See Foot (1985) and Nussbaum (1999)

virtue runs full and infallible virtue together. Full virtue is the disposition to act well; infallible virtue is perfect conformity with the demands of morality.<sup>43</sup> The virtue theorist is uniquely well-positioned to make this distinction. Now the task is to show how this distinction might shed light on excuses like the ones afforded to Erdemović and Vican.

## 5. Explaining Excuse

Suppose that Vican and Erdemović are fully virtuous in my sense. Of course, I do not take this starting assumption to be obviously true, but I make it in order to put the account in view; I will relax the claim in due course. If they are fully virtuous, they are disposed to act permissibly, and accordingly they have the beliefs, desires, and other mental states that ground the virtuous disposition. Vican and Erdemović act impermissibly in their respective circumstances, but our folk psychological judgment is that their failure is compatible with their continued possession of a virtuous disposition. To use the notion described before, they are fully but fallibly virtuous. In general, agents who are fully virtuous but act impermissibly are not fully blameworthy for what they do. So Vican and Erdemović are excused.

The preceding paragraph sketches the explanation I hope to offer for cases of excuse by duress. As I stated at the outset, the first part of the explanation needs to be relaxed. It was not part of their description that Erdemović and Vican are fully virtuous people in my sense. So why think that

---

<sup>43</sup> A technical caveat is needed. It may be reasonable to attribute exactly one disposition for impermissible behaviour to those who are fully virtuous. This disposition is unfortunately hard to state succinctly. It is *the disposition to act impermissibly when subjected to conditions such that impermissible behaviour is compatible with the continued possession of the relevant virtuous disposition*. This follows from the definition of full virtue. Unlike fully and infallibly virtuous people, those who are fully but fallibly virtuous may act impermissibly when (and only when) they are subjected to conditions of the sort described. So, it may be appropriate to say that those who are fully but fallibly virtuous have this peculiar disposition for impermissible conduct. Technically, then, the fully virtuous person is someone to whom it is inappropriate to attribute any disposition for impermissible conduct *except for the one disposition just described*. Thanks to an anonymous referee for making this observation.

the explanation above is applicable? Although it is unlikely that they are fully virtuous in all respects, I think it is plausible that these agents are fully virtuous *in the relevant respect*. If there is more than one virtue, each one will have a different domain of relevance. When someone performs a wrongdoing, they skirt a requirement that will be in the domain of one of the identified virtues. If someone is excused for their wrongdoing, they must be fully virtuous in the relevant way, i.e. they must have the relevant virtue.

Is it clear that Vican and Erdemović are fully virtuous in the relevant respect? That is, do they have the virtue that relates to the kind of requirement they skirt? Plausibly, both agents have the virtue that W. D. Ross called *nonmaleficence*, i.e. the disposition to refrain from injuring others.<sup>44</sup> One can find evidence for their possession of this virtue in their pacifist and anti-war activities prior to their wrongful acts. Erdemović's pacifism was apparently genuine.<sup>45</sup> Vican took measures to resist the Nazi regime, including breaking a contract that would have required him to harm others. Both of these agents experience psychological turmoil when they injure others *because* they are disposed to refrain from this kind of behaviour.

One might still worry that Vican and Erdemović do not actually have the virtue of nonmaleficence for other reasons. There are people who would give their lives in Vican and Erdemović's respective circumstances, and it might seem like *those* people are properly called virtuous in this respect. But recall that I distinguish full and infallible virtue. On this picture, the people who would give their lives, even in terrible circumstances like those facing Vican and Erdemović, are fully *and* infallibly virtuous. They have a psychological composition such that they are not only disposed to act well, but they also always manifest that disposition (barring ignorance about matters of fact). This psychological composition is the moral ideal. There is no doubt that

---

<sup>44</sup> See Ross (1930:21-22)

<sup>45</sup> Some testified during Erdemović's trial that he "hated the war and the army but believed that he simply had to do all of it; he was not the sort of person to kill of his own free-will" (*Prosecutor v. Erdemović*, 1998).

Vican and Erdemović fall short of this ideal, understood as full *and* infallible virtue, but it is not clear that they fall short in the technical sense of full virtue. In full, the virtue solution has three parts:

*Virtue solution*

1. *Definition:* E is fully virtuous in respect *c* if and only if E is disposed to act permissibly in all *c*-cases.

2. *Psychological Claim:* Being disposed to act permissibly in all *c*-cases is compatible with acting impermissibly in a subset of *c*-cases. Call these tolerable *c*-cases.

3. *Sufficiency Claim:* If E is fully virtuous in respect *c*, but acts impermissibly in a tolerable *c*-case, E is at least partially excused.

The first claim is just the definition of full virtue. The psychological claim was defended in the previous section. The sufficiency claim says that possession of full virtue in a particular respect can excuse wrongdoings related to that virtue. The sufficiency claim will be discussed in this section.

Does the virtue solution make it too hard to deserve an excuse? Few people are saints. The average person is probably disposed to be deceptive, ungrateful, rude, dishonest, or mean-spirited, at least to some extent. If this is true, wouldn't the virtue solution imply that only very exceptional people deserve the excuse of duress? I think not. Of course, I agree that the average person is not fully honest, grateful, kind, or courageous. It would be foolish to say otherwise. But no cases of genuine duress appear to involve these virtues. Take the transgression of breaking a promise. If someone breaks a promise and deserves the excuse of duress, it must be true that (a) they are required to keep the promise, (b) they are facing a considerable threat, and (c) they deserve reduced blame. It

seems to me that there are no obvious cases of this sort. If the person is threatened with corporal punishment or death unless they break their promise, then it is usually the case that breaking the promise is permissible. In such cases, there is no need for an excuse. If the person is threatened with something trivial, e.g. an indignant glance from a stranger, and breaks the promise, then it is usually the case that breaking the promise is impermissible. In such cases, there are usually no grounds for excuse.

My point is that the excuse of duress resists applicability in cases of promise-keeping and the other virtues listed above. If breaking the promise, telling the lie, or making the rude remark is impermissible, there are usually no grounds for excuse; and if breaking the promise, telling the lie, or making the rude remark is permissible, there is usually no need for an excuse in the strict sense. The crucial question is this: how many people possess full virtue in the particular respect that is required for the excuse of duress? I tentatively identified the relevant virtue as Rossian nonmaleficence, i.e. the disposition to refrain from injuring others. Though I can do little more than speculate, I would venture that the incidence of full nonmaleficence is greater than the incidence of full honesty, courage, or gratitude. At least, it seems plausible that the incidence of nonmaleficence might more closely match the number of people who deserve excuses for the kinds of crimes committed by Vican and Erdemović. This is especially plausible when one bears in mind that full nonmaleficence requires only the disposition to refrain from injuring others and does *not* require perfect conformity with morality.

Still, some might worry that the virtue solution is an overcorrection of the error in the Humean view. In Sher's counterexample, Alphonse is disposed to act permissibly in almost every circumstance, but he is disposed to act impermissibly in the atypical case that he makes a mistake. If Alphonse makes a mistake on one occasion and behaves impermissibly, he is blameworthy for his act, even though he is disposed to act well in other contexts. Since Alphonse is excused on the Humean view, Sher concludes that the Humean view fails. The virtue solution avoids this



counterexample by insisting on the standard of full virtue. Since Alphonse is disposed to act cruelly when he makes a mistake, he is presumably lacking in full virtue in that respect, and so he is blameworthy for his act. The potential problem cases for the virtue solution run in the opposite direction. The view is potentially too conservative; it precludes excuses for people who deviate from full nonmaleficence in any regard. This objection seems to concern an *inversed-Alphonse* case with respect to nonmaleficence. In an *inversed-Alphonse* case, the agent is disposed to refrain from harming other people in his current circumstances but, unlike Vican and Erdemović, he is disposed to act violently in at least one other nontrivial context.<sup>46</sup> If *inversed-Alphonse* behaves like Vican or Erdemović under threat, he will not deserve an excuse on my proposal. This may seem to render the virtue solution just as unappealing as the Humean view.

It is harder than it may seem to devise an *inversed-Alphonse* case in which the agent clearly deserves an excuse. There are two important constraints pushing against it. To see this, consider two *inversed-Alphonse* cases:

(A) Josef shares Vican's predicament. Like Vican, Josef is disposed to refrain from harming his fellow prisoners but he punishes them under threat from guards. But, unlike Vican, Josef is an abusive parent. If he feels insecure, angry, dissatisfied, or otherwise irritable, Josef is disposed to abuse his children in ways that are not dissimilar from the abusive acts that he is ordered to perform against his fellow prisoners.

(B) Lukas too shares Vican's predicament. Unlike Vican, Lukas hates all people except his fellow Austrians. As it happens, all of the prisoners under Lukas'

---

<sup>46</sup> I say 'nontrivial' context because the agent in question must be disposed to act violently in a context that is serious enough to preclude his possession of full nonmaleficence.

supervision are Austrian, so Lukas is disposed to refrain from harming them.

Nonetheless, Lukas punishes his supervisees under threat from the guards.

In both cases, the agent deviates from full nonmaleficence and acts impermissibly under threat. The deviation is different in each instance. In (A), Josef deviates from full nonmaleficence because he is disposed to abuse his children. In (B), Lukas deviates from full nonmaleficence because he hates non-Austrians and thereby has the broad disposition to harm them. Intuitions about these cases may vary. But there are principled reasons to think that neither Josef nor Lukas are worthy of moral excuse.

Josef violates what I call the *injury constraint*. The injury constraint says that agents may be excused for wrongdoing only if they are morally injured by their behaviour. When an act is morally injurious, it requires that the agent "transgress deeply held moral beliefs and expectations" in ways that are deleterious to their emotional, psychological, and social health (Litz *et. al.* 2009: 695). Moral injuries are usually accompanied by intense feelings of guilt, shame, regret, and alienation after the transgression of one's moral boundaries.<sup>47</sup> Although the extensive study of moral injury is still in its infancy, the concept has been familiar for some time. Over two decades ago, Jonathan Shay introduced the term to refer to the psychological and social harm caused by the betrayal of a person's sense of right and wrong.<sup>48</sup> Philosophers may notice a rough but perceptible relationship between moral injury and unthinkability in Bernard Williams' sense of the latter term. If pushed to perform an unthinkable act, an agent doesn't simply choose to do something that she doesn't want to do; rather, she is pushed to take an option that she did not previously even countenance *as an option*.<sup>49</sup> Moral injury could be understood as the harm done to someone when their conduct is (or was) unthinkable

---

<sup>47</sup> See Linz, *et. al.* (2009: 698-99)

<sup>48</sup> For his most thorough treatment of moral injury, see Shay (1995). See also Shay (2014).

<sup>49</sup> See Williams (1981b: 128-129).

by their own lights. If Williams is right, and the performance of an unthinkable act is harmful to the agent, then unthinkability may help explicate the harm involved in moral injury.

There is some sense in which people who act under duress are themselves victimized or harmed by their behaviour. Obviously, people like Vican and Erdemović are not harmed in the same sense that their victims are harmed by their actions. The harm done to Vican or Erdemović is not corporeal. Plausibly, their injury is moral; they are harmed by the transgression of their own moral boundaries. Although there is limited information about Vican's specific case, Primo Levi—an Italian survivor of Auschwitz—describes the suffering of those who occupied the ‘gray zone’ between victim and perpetrator. He recalls an illuminating remark from someone in Vican's position: “Certainly, I could have killed myself or got myself killed; but I wanted to survive, to avenge myself and bear witness. You mustn't think that we are monsters; we are the same as you, only much more unhappy” (Levi 1988: 41). With this in mind, Levi concludes that:

Conceiving and organizing the squads was National Socialism's most demonic crime. Behind the pragmatic aspect (to economize on able men, to impose on others the most atrocious tasks) other more subtle aspects can be perceived. This institution represented an attempt to shift onto others specifically, the victims—the burden of guilt, so that they were deprived of even the solace of innocence. (Levi 1988: 41)

There is evidence that Erdemović too suffered moral injury. According to trial documents, he experienced such severe post-traumatic stress that he was initially unable to stand trial by order of psychiatric authorities.<sup>50</sup> During his sentencing, judges in the trial chamber noted that “the post-

---

<sup>50</sup> See *Prosecutor v. Erdemović* (1998). It should be noted that post-traumatic stress disorder and moral injury are distinct on most models (see Linz *et al.* 2009: 696). Nonetheless, the conditions are related (Linz, *et. al.* 2009: 697) and may not have been distinguished by the clinicians treating Erdemović.

traumatic stress which [Erdemović] suffered...in the aftermath of the...atrocities demonstrates how he himself has suffered from being forced to commit the killings against his will" (*Prosecutor v. Erdemović*, 1998). The judges acknowledged his "reluctance to participate and his reaction to having to perform this gruesome task" (*Prosecutor v. Erdemović*, 1998).

The injury constraint explains why Josef doesn't obviously deserve an excuse for acting wrongly in (A). Given his abusive disposition, it is unlikely that violence is unthinkable for Josef, and accordingly, it is hard to say he suffers moral injury of the sort described earlier. Of course, this is an empirical claim about Josef's psychology, so it is open to refutation on empirical grounds. But, as long as this speculation is plausible, Josef's case is no obstacle for the virtue solution.

There is a reading on which the agent in (B), Lukas, does not run afoul of the injury constraint. Given his good will toward fellow Austrians, Lukas might struggle immensely to punish those Austrians under his supervision. Indeed, it is plausible that Lukas thinks violence against fellow Austrians is unthinkable in Williams' sense.<sup>51</sup> But Lukas manages to meet the injury constraint by instead violating a second constraint. According to the *praiseworthiness constraint*, an excused agent's reluctance to act must be itself morally praiseworthy. In other words, it is inappropriate to excuse someone for acting wrongly when they would not be praiseworthy if they were to act rightly in the same context. No one wants to excuse an agent who finds it difficult to injure or kill other people for morally objectionable or misguided reasons. Recall Erdemović's sadistic comrade, who finds it very difficult to murder the civilians, but only because he prefers to treat them even more heinously. The praiseworthiness constraint explains why the sadistic comrade is not entitled to an excuse. He is reluctant to act, but his reluctance is grounded in something morally objectionable, so he doesn't deserve an excuse.

In (B), Lukas may struggle considerably to comply with the guard's orders, but his experience

---

<sup>51</sup> Thanks to an anonymous referee for highlighting this case.

of struggle is not itself praiseworthy. Lukas resembles Nomy Arpaly's extremist, who doesn't believe that murder is generally immoral, but thinks it is wrong to kill fellow Jews.<sup>52</sup> Arpaly says that the Jewish extremist would not deserve praise if he were to resist injuring or otherwise harming a fellow Jewish person. The extremist chooses peaceful inaction because his potential victim is Jewish, but someone's ethnicity or religion has no bearing on why they should not be killed, so the extremist's act appears accidentally right. If he were to act rightly, Lukas would resist because his fellow prisoners are Austrian, but someone's nationality has no bearing on why they should not be attacked.<sup>53</sup> Like the extremist, Lukas would not deserve praise if he were to act rightly because the grounds for his resistance would be unconnected to reasons why he ought to resist. His act would be accidentally right (in the way that normally seems to preclude praiseworthiness).<sup>54</sup>

Neither Josef nor Lukas are clear counterexamples to the virtue solution. If the agent is prone to violence in a way that precludes full nonmaleficence, then there are reasonable grounds to bar them from an excuse for any impermissible act of violence. In fact, I think there is *no* inverted-Alphonse case that includes an agent who is obviously excusable. There are two layers of difficulty standing in the way of developing this example.

First, the agent in the chosen example must deviate from the particular virtue of nonmaleficence, as in (A) and (B) above. This means that my opponents cannot appeal to people who are disposed to behave violently when they are involuntarily drugged, intoxicated, or otherwise

---

<sup>52</sup> Arpaly 2002: 227

<sup>53</sup> I mean here that the rationalizing explanation of why Lukas resists injuring his fellow Austrians (i.e. his 'motivating reason' for resisting) is that *the prisoners under his supervision are fellow Austrians*. The fact that rationalizes someone's behaviour needn't appear as an explicit or occurrent thought in the agent's mind. This is why Huckleberry Finn can free his friend, Jim, for the right reasons, even though he explicitly thinks that his act is morally wrong. See Arpaly 2002 and Markovits 2010.

<sup>54</sup> I will not take a stand here on the nature of this accidentality, nor do I want to advance any particular account of moral worth. The crucial point is that Lukas would not deserve praise if he were to act rightly, just as the extremist does not deserve praise for acting rightly in Arpaly's example.

incapacitated, since these cases do not clearly count against someone's possession of the virtue.<sup>55</sup> Nor can they appeal to people whose moral flaws are best understood as violations of beneficence rather than nonmaleficence, like those who are merely impolite, mean, or self-centered.

Second, the agent in the chosen example must meet the injury and praiseworthiness constraints, which seem jointly incompatible with deviations from full nonmaleficence. There is a general argument for this claim. Suppose that *S* deviates from full nonmaleficence and acts wrongly in circumstances similar to those facing Vican. If *S* meets the injury constraint, then (i) there is some difference between her current situation and the situations in which she is violent, and (ii) *S* thinks this difference is morally significant. In Lukas' case, the difference is the victim's nationality; he is disposed to violence when the victims are non-Austrian, but he is uninterested in hurting any Austrians. If *S* deviates from full nonmaleficence, and meets the injury constraint, there must be *something* on which her selective violence rests. If there were no such conceptual division—that is, if her disposition to violence was entirely unprincipled—then it is hard to see how *S* could plausibly find violence thinkable in one context but find it unthinkable in other contexts. There must be some key difference that *S* believes is morally salient—say, the victim's race, nationality, occupation, weight, fashion, or relation to her. But, if this is true, then *S* inevitably violates the praiseworthiness constraint. Her belief in the moral salience of some feature (e.g. race, nationality, etc.) will appear in the best explanation of why *S* struggles to act violently under threat. If someone's disposition to violence rests on the presence or absence of some feature, then it is natural to rationalize their struggle to act violently by appealing to their belief about the moral salience of that feature. To rationalize why the Jewish extremist refrains from violence against his fellow Jews, Arpaly appeals

---

<sup>55</sup> These people are akin to those who are disposed to violence if an evil neuroscientist interfered with their brain and triggered violent reactions. This doesn't tend to count against possession of full nonmaleficence. Voluntary incapacitation cases are more complicated. Suppose I know that I am violent when intoxicated. If I want to hurt someone I dislike, I might drug myself before visiting them. This does seem incompatible with full nonmaleficence. I have a violent disposition, even though it mediated by the use of drugs.

to the extremist's belief that Jewish people, and only Jewish people, should not be harmed. Any rationalizing explanation of the extremist's action that didn't include this belief would seem woefully inaccurate. Now it should be clear why *S* must violate the praiseworthiness constraint. Given that *S* deviates from full nonmaleficence, her beliefs about moral salience are bound to be misguided or objectionable. After all, her belief implies that some decent people deserve to be treated violently. This is because someone who deviates from full nonmaleficence must be disposed to behave violently in some unjustified cases, i.e. some cases involving decent people. (If she was disposed to behave violently only in *justified* cases, then *S* does not deviate from full nonmaleficence.) Inevitably, *S*'s resistance will be grounded in something morally misguided or objectionable. So, in order to meet the injury constraint, *S* must run afoul of the praiseworthiness constraint. In conjunction with the first layer of difficulty, the injury and praiseworthiness constraints ensure that the prospects for a counterexample to the virtue solution are dimmer than they may first appear.

I have tried to show that the virtue solution is neither too strict nor too liberal. It means that someone who merely feels psychological distress or someone who feels no distress whatsoever does not deserve an excuse for acting wrongly under duress. At the same time, it does not require that the accused is virtuous in all respects nor does it require that she never injures other people. It simply requires that the accused possesses full nonmaleficence, i.e. the disposition to refrain from injuring others.

## **6. Norm-splitting revisited**

It may appear that norm-splitting reemerges on my account. After all, there are two normative standards on my view: full virtue and infallible virtue. A 'split' in standards apparently explains the excuse afforded to Erdemović and Vican. This suggests that the virtue solution is not really a departure from norm-splitting, but rather just another version of it.

The sense in which my view 'splits' norms is endemic to the concept of excuse. As I use the term, when one excuses an action, one judges that the conduct is impermissible but not blameworthy. As a result, any explanation of excuse must make sense of how *impermissibility* and *blameworthiness* come apart. This means that accepting the concept of excuse just means that there are standards for blameworthiness that come apart from standards of conduct. In this sense, my view does 'split' norms.

But norm-splitting as I use the term is a more specific proposal. The norm-splitting approach appeals specifically to *deontic* norms in order to explain how blameworthiness and impermissibility come apart. This view can seem appealing. The concept of excuse already has the concept of impermissible conduct built into it, so it is natural (albeit misguided) to appeal to more deontic norms, like those governing motivation or character, in order to understand excuses. So, while my view 'splits' norms in sense endemic to the explanation of excuse, it is not norm-splitting as I use the term.

There is another way to show how my proposal diverges from the norm-splitting approach. Any norm-splitting approach involves some kind of brute fragmentation among norms. On Rosen's proposal, one is required to perform an act, but not required to be moved to do it.<sup>56</sup> On Hume's proposal, one is required to perform an act, but not required to be generally disposed to perform acts of that type. In both cases, one must take the fragmentation as a brute fact about morality. But these facts do seem to require some explication, especially if they are expected to help us understand other concepts like excuse.

In contrast, on the virtue solution, morality speaks univocally: if one is required to perform an action, one is required to be moved to perform that action, and one is required to be disposed to perform all actions of that sort. In other words, the deontic norms governing action, motivation, and

---

<sup>56</sup> Rosen (2014: 86) admits that his version of norm-splitting calls out for explanation, but doesn't offer one in his essay.



character line up. Vican and Erdemović are required to resist their orders, required to be moved to resist their orders, and required to be disposed to resist maleficent orders in general. No schism among deontic norms is needed to make the virtue solution work.

Moreover, the view needn't treat the distinction between full and infallible virtue as a brute fact about morality. There is an available non-moral explanation of why the standards for full virtue and infallible virtue come apart: agential dispositions tolerate failures to manifest. This fact about how we attribute dispositions makes it possible to distinguish full and infallible virtue. If our folk psychological practices were different, and dispositions did not tolerate counterinstances, there would be no way to conceptualize the two standards used in the virtue solution. The possibility of this kind of explanation is one key difference between the norm-splitting approach and the approach defended in this essay.

Although the virtue solution appeals to a folk psychological truth, it is clearly superior to the purely psychological explanation of excuse by duress. The purely psychological view went wrong in grounding excuses in the experience of inner struggle directly. This meant that anyone who struggles to do the wrong thing will be excused, including those who struggle because they want to do something even worse or value trivial things too highly. I have suggested that an excuse by duress is not grounded in the experience of inner struggle *itself*, but rather in the fact that inner struggle *reveals*, namely that the agent in question is disposed to do the right thing.

## 7. Conclusion

The excuse afforded to Vican and Erdemović rests in the fact that they are nonmaleficent, or disposed to refrain from injuring others, even in terrible circumstances like theirs. However, *pace* Humean norm-splitters, they do not count as nonmaleficent because virtuous dispositions issue in

occasional impermissible conduct.<sup>57</sup> Rather, they count as nonmaleficent because virtuous dispositions tolerate failures to manifest, which is a fact about the folk practice of ascribing any disposition to act. This is the proposed explanation of the excuse afforded to Vican and Erdemović.

There are remaining questions about the extent of the exculpation offered by the excuse of duress. First, does duress ever offer *full* exculpation? Conditions like blameless ignorance of fact and blameless incompetence seem to warrant full exculpation when applicable. Conditions like provocation and entrapment appear to warrant only partial exculpation when they apply.<sup>58</sup> It is notable that those who treat duress as an excuse tend to group duress with merely mitigating conditions like provocation.<sup>59</sup> The virtue solution can explain why duress does not seem to fully excuse wrongful actions. When someone pleads blameless ignorance or incompetence, they do not typically admit moral deficiency, though they may admit to epistemic, physical, or technical failings. In contrast, someone who pleads duress or provocation *does* admit some kind of moral deficiency. To use a familiar phrase in English law, the excuse afforded by duress and provocation are concessions to human frailty.<sup>60</sup> They apply when a person has “perhaps to some exceptionally limited extent, a ‘guilty mind’ [but] whose mind is less guilty than is his who acts as he does but under no such constraint” (*R. v Howe*, 1987). The virtue solution gives us a way to articulate this limited deficiency: these agents are fully but fallibly virtuous in the relevant respect. The moral failure embodied in full but fallible virtue is worthy of blame, but not completely so.

---

<sup>57</sup> Here I contrast Humean claim that virtuous dispositions *issue in* bad conduct with my claim that virtuous disposition *tolerate* bad conduct. The Humean claim is that someone can still count as virtuous even if they are disposed to act impermissibly to some extent. I said that virtuous dispositions (psychologically) tolerate bad conduct; a fully virtuous person cannot be disposed to act impermissibly, but she can act impermissibly on an occasion.

<sup>58</sup> For discussion of provocation as a partial excuse, see Dressler (1988). and for entrapment as a partial excuse, see Hastings (1986). Both of these accounts apply to uses of these terms in legal contexts, but they can be applied in the moral context as well.

<sup>59</sup> See, for example, Reed (1996) and Uniacke (2007).

<sup>60</sup> This phrase is regularly used in the context of the provocation defense in English law since (at least) *R v Hayward*, 1833. Lord Hailsham uses this phrase to describe duress in *R v. Howe*, 1987.

Second, does duress partially exculpate to different degrees? It is important to be mindful about exactly what it meant by this question. Duress itself is scalar. It is obvious that threats can range from severe to mild. The former occurs when threatened with death and the latter happens when merely threatened with a glance of disapproval. But it doesn't follow that duress *exculpates* to different degrees. The exculpation offered by duress does not come in degrees. Once the severity of a threat passes a threshold, the agent is entitled to the excuse of duress and the mitigation it provides. That is, if the threat is serious enough to warrant excuse, it doesn't matter exactly how bad the threat happens to be. Suppose that the threat facing Erdemović were more devastating. For example, suppose his superiors threaten to kill him and also threaten to forcibly enlist his son to replace him on the killing squad. Perhaps this additional threat still doesn't suffice to make Erdemović's conduct permissible (we may assume) but it nonetheless makes the threat more devastating and serious for him. I don't think Erdemović would be entitled to *more* exculpation in this scenario, at least not more than he was already guaranteed by the threat against his life. The threats could be slightly intensified. Perhaps Erdemović's superiors promise to forcibly recruit his son *and* wife to serve in the military. He doesn't seem worthy of more exculpation than he already is, assuming again that the increased threats do not justify his behaviour.

The virtue solution preserves this feature of duress. I said earlier that nonmaleficence in the sense of full virtue can come in degrees. It is possible for one's disposition to refrain from harming others to match the relevant virtuous disposition to a greater or lesser extent. But possession of full nonmaleficence is bivalent. That is, someone is either fully nonmaleficent or she is not. When someone is fully nonmaleficent, but acts wrongly, she deserves the partial exculpation that the excuse of duress provides. When she deviates from full nonmaleficence, and acts wrongly, she is not entitled to the same excuse.

I have claimed that the excuse of duress offers only partial exculpation for wrongdoing and the degree of this exculpation doesn't vary with the severity of the threat after a threshold of severity

is met. So, exactly how much blame do people like Vican and Erdemović deserve? Even if there were general rules about the warranted degree of mitigation, it would be hard to describe their degree of blameworthiness in non-comparative terms.<sup>61</sup> They deserve less blame for acting wrongly than their complicit, enthusiastic, or sadistic counterparts. This essay is not an attempt to make their level of blameworthiness any more precise than this initial observation. My goal is more modest. I merely tried to lay a theoretical foundation for the excuse of duress.

The proposed explanation is not purely normative nor is it purely psychological. It doesn't propose brute discrepancies in deontic norms, nor does it say that mere psychological resistance is sufficient for excuse. It coheres with the commonsense view: if I am required to perform an act, I am required to be moved to perform it, and I am required to have a general disposition that includes the doing it. The virtue solution carves out room for the excuse of duress using an overlooked distinction between full and infallible virtue. This distinction gives us the resources to understand duress, and perhaps partial excuse generally, without attributing any inexplicable doubletalk to morality.

---

<sup>61</sup> Wasik (1982: 529)

## References

- Annas, Julia. 2004. "Being Virtuous and Doing the Right Thing." *Proceedings and Addresses of the American Philosophical Association* 78 (2): 61–75.
- . 2005. "Comments on John Doris's Lack of Character." *Philosophy and Phenomenological Research* 71 (3): 636–42.
- . 2011. *Intelligent Virtue*. OUP Oxford.
- Armstrong, David M. 1961. *Perception And The Physical World*. Humanities Press.
- Arpaly, Nomy. 2002. "Moral Worth." *Journal of Philosophy* 99 (5): 223–45.
- . 2003. *Unprincipled Virtue: An Inquiry Into Moral Agency*. Oxford University Press.
- Ashwell, Lauren. 2010. "Superficial Dispositionalism." *Australasian Journal of Philosophy* 88 (4): 635–53.
- . 2015. "Conflicts of Desire: Dispositions and the Metaphysics of Mind." In *Putting Powers to Work*, edited by Jonathan Jacobs. Oxford: Oxford University Press.
- Austin, John. 1956. "A Plea for Excuses." *Proceedings of the Aristotelian Society* 57: 1–30.
- Baron, Marcia. 2006. "Excuses, Excuses." *Criminal Law and Philosophy* 1 (1): 21–39.
- . 2014. "II—Culpability, Excuse, and the 'Ill Will' Condition." *Aristotelian Society Supplementary Volume* 88 (1): 91–109.
- Bayles, Michael D. 1982. "Character, Purpose, and Criminal Responsibility." *Law and Philosophy* 1 (1): 5–20.
- Bennett, Jonathan. 1979. "Accountability." In *Philosophical Subjects: Essays Presented to P.F. Strawson*, edited by ZV Straaten. Oxford: Oxford University Press.
- Bird, Alexander. 1998. "Dispositions and Antidotes." *Philosophical Quarterly* 48 (191): 227–34.
- Bird, A., and T. Handfield. 2008. "Dispositions, Rules and Finks." *Philosophical Studies* 140 (2): 285–98.
- Brandt, R. B. 1958. "Blameworthiness and Obligation." In *Essays in Moral Philosophy*, edited by A. I. Melden. University of Washington Press.
- . 1970. "Traits of Character: A Conceptual Analysis." *American Philosophical Quarterly* 7 (1): 23–37.
- Brooks, Rosa E. 2012. "Law in the Heart of Darkness: Atrocity and Duress." *Virginia Journal of International Law* 43.

- Choi, Sungho. 2012. "What Is a Dispositional Masker?" *Mind* 120 (480): 1159–71.
- . 2009. "The Conditional Analysis of Dispositions and the Intrinsic Dispositions Thesis." *Philosophy and Phenomenological Research* 78 (3): 568–90.
- Clark, Andy & Chalmers, D. 1998. "The extended mind." *Analysis* 58(1): 7-19.
- Clarke, Randolph. 2008. "Intrinsic Finks." *Philosophical Quarterly* 58 (232): 512–18.
- . 2010. "Opposing Powers." *Philosophical Studies* 149 (2): 153–60.
- Doris, John M. 2002. *Lack of Character: Personality and Moral Behavior*. Cambridge University Press.
- Dressler, J. 1988. "Provocation: Partial Justification or Partial Excuse?" *The Modern Law Review*, 51(4), 467-480.
- Duff, Antony. 2007. *Answering for Crime: Responsibility and Liability in the Criminal Law*. Hart Pub.
- . 2006. "Excuses, Moral and Legal: A Comment on Marcia Baron's 'Excuses, Excuses.'" *Criminal Law and Philosophy* 1 (1): 49–55.
- Fara, Michael. 2005. "Dispositions and Habituals." *Noûs* 39 (1): 43–82.
- . 2008. "Masked Abilities and Compatibilism." *Mind* 117 (468): 843–65.
- Finkelstein, Claire. 2002. "Excuses and Dispositions in Criminal Law." *Buffalo Criminal Law Review* 6 (1): 317–59.
- Foot, Philippa. 1983). "Utilitarianism and the Virtues." *Proceedings and Addresses of the American Philosophical Association*, 57(2), 273-283
- Frankfurt, Harry G. 1969. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66 (3): 829–39.
- Gardner, John. 2007. *Offences and Defenses: Selected Essays in the Philosophy of Criminal Law*. Oxford University Press.
- Handfield, Toby. 2008. "Unfinkable Dispositions." *Synthese* 160 (2): 297–308.
- Hastings, K.A., 1986. "Entrapment and Denial of the Crime: A Defense of the Inconsistency Rule." *Duke Law Journal*, 1986(5), pp.866-891.
- Heil, John, and C. B. Martin. 1998. "Rules and Powers." *Philosophical Perspectives* 12 (S12): 283–312.
- Heim, Sarah. 2013. "The Applicability of the Duress Defense to the Killing of Innocent Persons by

- Civilians.” *Cornell International Law Journal* 46 (1): 165–90.
- Herman, Barbara. 1981. “On the Value of Acting From the Motive of Duty.” *Philosophical Review* 90 (3): 359–82.
- Hieronymi, Pamela. 2007. “Rational Capacity as a Condition on Blame.” *Philosophical Books* 48 (2): 109–23.
- Hilton, Fern Overbey. 2004. *The Dachau Defendants: Life Stories from Testimony and Documents of the War Crimes Prosecutions*. McFarland & Company Publishers.
- Hobart, R. E. 1934. “Free Will as Involving Determination and Inconceivable Without It.” *Mind* 43 (169): 1–27.
- Hume, David. 1738. *A Treatise of Human Nature*. Oxford University Press.
- Husak, Douglas. 2005. “On the Supposed Priority of Justification to Excuse.” *Law and Philosophy* 24 (6): 557–94.
- Johnston, Mark. 1992. “How to Speak of the Colors.” *Philosophical Studies* 68 (3): 221–63.
- Kahan, Dan & Nussbaum, Martha. 1996. "Two Conceptions of Emotion in Criminal Law." *Columbia Law Review* 96 (2).
- Kamtekar, Rachana. 2004. “Situationism and Virtue Ethics on the Content of Our Character.” *Ethics* 114 (3): 458–91.
- Kelly, Erin. 2012. “What Is an Excuse?” In *Blame: Its Nature and Norms*, edited by Neal Tognazzini and D. Justin Coates. Oxford; New York: Oxford University Press.
- Kent, Bonnie. 2012. “Dispositions and Moral Fallibility: The UnAristotelian Aquinas.” *History of Philosophy Quarterly* 29 (2).
- Kripke, Saul A. 1982. *Wittgenstein on Rules and Private Language*. Harvard University Press.
- Levi, Primo. 1988. *The Drowned and the Saved*. Translated from the Italian by Raymond Rosenthal. Summit Books.
- Lewis, David. 1997. “Finkish Dispositions.” *Philosophical Quarterly* 47 (187): 143–58.
- Linz, B. Stein, N. Delaney, E. Lebowitz, L. Nash, W. Silva, C. Maguen, D. 2009. "Moral injury and moral repair in war veterans: A preliminary model and intervention strategy." *Clinical Psychology Review* 29(8): 695-706.
- Manley, David, & Ryan Wasserman. 2008. “On Linking Dispositions and Conditionals.” *Mind* 117 (465): 59–84.
- Markovits, Julia. 2010. “Acting for the Right Reasons.” *Philosophical Review* 119 (2): 201–42.

- Martin, C. B. 1994. "Dispositions and Conditionals." *Philosophical Quarterly* 44 (174): 1–8.
- McKenna, Michael. 2012. *Conversation & Responsibility*. OUP.
- Moore, Michael S. 1997. *Placing Blame: A Theory of the Criminal Law*. Oxford University Press.
- Murphy, Jeffrie G., and Jean Hampton. 1990. *Forgiveness and Mercy*. Cambridge University Press.
- Murphy, J. G. 1992. *Retribution Reconsidered: More Essays in the Philosophy of Law*. 1992 edition. Dordrecht Holland ; Boston: Springer.
- Nussbaum, Martha C. 1999. "Virtue Ethics: A Misleading Category?" *The Journal of Ethics*, vol. 3, no. 3: 163–201.
- Prosecutor v. Erdemović, Sentencing Judgment. 1998. Case No. IT-96-22tbis. United Nations Criminal Tribunal for the Former Yugoslavia.
- Raz, Joseph. 1975. "Permissions and Supererogation." *American Philosophical Quarterly* 12 (2): 161–68.
- Reed, Alan. 1996. "Duress and Provocation as Excuses to Murder: Salutory Lessons from Recent Anglo-American Jurisprudence." *Journal of Transnational Law and Policy* 6(1): 51-92.
- R. v. Hayward. 1833. 6 C & P 157 172 ER 1188.
- R. v. Howe and Bannister, Judgement of Lord Hailsham. 1987. UKHL 2 WLR 568
- Richards, Norvin. 1987. "Acting Under Duress." *Philosophical Quarterly* 37 (146): 21–36.
- Rosen, Gideon. 2002. "Culpability and Ignorance." *Proceedings of the Aristotelian Society* 103 (1): 61–84.
- Ross, W. D. 2002. *The Right and the Good*. Clarendon Press.
- Russell, Daniel C. 2009. *Practical Intelligence and the Virtues*. Oxford University Press.
- Scanlon, Thomas. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Belknap Press of Harvard University Press.
- Setiya, Kieran. 2012. *Internal Reasons*. In Kieran Setiya & Hille Paakkunainen (eds.), *Internal Reasons: Contemporary Readings*. MIT Press.
- Sher, George. 2007. *In Praise of Blame*. Oxford: Oxford University Press.
- . 2012. "Wrongdoing and Relationships: The Problem of the Stranger." In *Blame: Its Nature and Norms*, edited by D. Justin Coates and Neal A. Tognazzini, 49–65. Oxford University Press.
- Shay, Jonathan. 1995. *Achilles in Vietnam: Combat Trauma and the Undoing of Character*. New York: Simon & Schuster.



- . 2014. "Moral Injury." *Psychoanalytic Psychology* 31(2): 182-191.
- Slote, Michael A. 2001. *Morals From Motives*. Oxford University Press.
- Smart, J. J. C., and Bernard Williams. 2013. *Utilitarianism: For and Against*. Cambridge University Press.
- Smith, Angela M. 2008. "Character, Blameworthiness, and Blame: Comments on George Sher's in Praise of Blame." *Philosophical Studies* 137 (1): 31–39.
- Sreenivasan, Gopal. 2002. "Errors About Errors: Virtue Theory and Trait Attribution." *Mind* 111 (441): 47–68.
- Strawson, Peter F. 1962. "Freedom and Resentment." *Proceedings of the British Academy* 48: 1–25.
- Sverdlik, Steven. 2011. *Motive and Rightness*. Oxford ; New York: Oxford University Press.
- Swanton, Christine. 2003. *Virtue Ethics: A Pluralistic View*. Clarendon Press.
- Tannenbaum, Julie. 2015. "Mere Moral Failure." *Canadian Journal of Philosophy* 45 (1): 58–84.
- Turns, D. 1998. "The International Criminal Tribunal for the Former Yugoslavia: The Erdemović Case." *International and Comparative Law Quarterly* 47: 461–74.
- Uniacke, Suzanne. 2007. "Emotional Excuses." *Law and Philosophy* 26(1): 95-117.
- United States v. Vican, Review and Recommendations. 1947. Case No. 000-Flossenburg-3. General Military Government Court at Dachau.
- Urmson, J., 1958, "Saints and Heroes", in *Essays in Moral Philosophy*, A. Melden (ed.), Seattle: University of Washington Press.
- Vuoso, George. 1987. "Background, Responsibility, and Excuse." *The Yale Law Journal* 96 (7): 1661–86.
- Wallace, R. Jay. 1996. *Responsibility and the Moral Sentiments*. Harvard University Press.
- . 2006. *Normativity and the Will: Selected Papers on Moral Psychology and Practical Reason*. Oxford University Press.
- . 2011. "Dispassionate Opprobrium: On Blame and the Reactive Sentiments." In *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, edited by Rahul Kumar, Samuel Freeman, and R. Jay Wallace. Oxford ; New York: Oxford University Press.
- Wasik, M. 1982. "Partial Excuses in the Criminal Law." *The Modern Law Review*, 45: 516-533.
- Watson, Gary. 1977. "Skepticism About Weakness of Will." *Philosophical Review* 86 (3): 316–39.

Williams, B. 1981a. Internal and External Reasons. In *Moral Luck: Philosophical Papers 1973–1980* (pp. 101-113). Cambridge: Cambridge University Press.

Williams, B. 1981b. Practical Necessity. In *Moral Luck: Philosophical Papers 1973–1980* (pp. 124-131). Cambridge: Cambridge University Press.

Yee, Sienho. 1997. "The Erdemovic Sentencing Judgment: A Questionable Milestone for the International Criminal Tribunal for the Former Yugoslavia." *Georgia Journal of International and Comparative Law* 26: 2.

Zaibert, Leo. 2012. "On Forgiveness and the Deliberate Refusal to Punish: Reiterating the Differences." *Journal of Moral Philosophy* 9 (1): 103–13.