Mental illness, exemption & moral exclusion: the role of interpretative generosity

Anna Hartford & Dan J. Stein

Published online: 28 Jun 2024.

Submit your article to this journal ⧉

View related articles ⧉

View Crossmark data ⧉

Routledge
Taylor & Francis Group

Mental illness, exemption & moral exclusion: the role of interpretative generosity

Anna Hartford and Dan J. Stein

ABSTRACT

Exemption from blameworthiness is bound to implicit or explicit claims of diminished agency, or even non-agency. This poses a dilemma in navigating moral relationships affected by mental illness. While it is crucial for assessments of responsibility to be responsive to the significance of mental illness, must this responsiveness come at a cost to symmetrical moral relations? In this paper we argue, contra recent critiques, that Strawsonian accounts of responsibility are able to navigate this dilemma, and can accommodate significant exculpation on the basis of mental illness while maintaining symmetrical relations. We understand this to be part of the larger recognition that while we have certain fundamental entitlements within our moral relationships, it is also possible to be *over-entitled*, and to expect too much of others. Our account draws, in particular, on the moral significance of difficulty. Difficulty is inherently scalar, and in drawing on this explanatory framework we explicate a theory of both excuse and exemption by degree. We argue for a significant realm of excuse on the basis of mental illness that is fully compatible with symmetrical relations. However, we also acknowledge the limits of this accommodation on Strawsonian grounds, and argue that these limits are justified and appropriate.

1. Introduction

Our practice of blaming one another, and holding one another responsible, has a fundamentally double nature. In one sense it is a moral burden to be deemed blameworthy. From this vantage, the desire to exempt can seem like a generous impulse. In another sense, however, an assertion of blameworthiness, and its extent, is implicitly bound to an assertion of moral agency. In turn, the withholding of blameworthiness, or its diminishment, is often bound to implicit or explicit claims of diminished agency. On prominent Strawsonian accounts of responsibility, these assessments are intimately related to evaluations of membership within, and exclusion from, moral communities established by mutual interpersonal demands and expectations.

Various philosophers have expressed this double nature of blame, and its essential role in our moral relationships. Angela Smith writes that 'being held responsible is as much a

CONTACT Anna Hartford ✉ annahartford@gmail.com 🖅 Neuroscience Institute, University of Cape Town, Cape Town, 7935, South Africa

privilege as it is a burden. It signals that we are a full participant in the moral community', and in turn that 'to deny someone this status … is a serious matter' (2005, 269). T. M. Scanlon argues that withholding blame often 'involves an attitude of superiority toward the person in question (something like the attitude of a parent toward a very young child) and thus represents a failure to take that person seriously as a participant in the relationship' (2008, 168).

In P. F. Strawson's seminal paper *Freedom and Resentment* ([1962] 1993) he describes a theory of moral responsibility founded on 'the demand for the manifestation of a reasonable degree of good will or regard, on the part of others'. Call this (following Watson 1993) the *basic demand*, which turns on the bearing of agents towards one another. In some respects the mutual assertion of this demand (and the mutual recognition of its legitimacy) establishes our shared participation in the moral community: we take ourselves to be answerable to each other on these grounds, and appropriately blamed where we fall short without a valid excuse. This shared acknowledgement underwrites relations of moral reciprocity and equality, or what we will be calling symmetrical relations.

So where we take the basic demand *itself* to be misplaced – where we think someone is not answerable to these expectations, and not appropriately blamed for disappointing them – we necessarily abandon symmetrical relations. At an extreme, Strawson describes the 'objective stance', where we treat people not so much as moral participants, but rather (in his now infamous phrase) as individuals to be 'managed or handled or cured or trained' ([[1962] 1993 [1962] 1993, 52). This framing has been extremely influential, including for accounts that are not explicitly Strawsonian, and many existing theories of moral responsibility quickly dispatch with cases of mental illness by evoking exemption (without giving much thought to the costs of such exemption).

The double nature of blameworthiness poses a dilemma in navigating our moral relationships affected by mental illness. While it is of course crucial for assessments of moral responsibility to be responsive to the significance of mental illness (and to the particular significance of particular mental illnesses), must this responsiveness come at a cost to relationships of moral equality and reciprocity?

Various philosophers have interpreted this dilemma as posing a serious challenge to Strawsonian accounts of responsibility. This challenge has been expressed both descriptively and normatively.

With regards to the descriptive challenge: some have argued that such accounts are inadequate explanations of our actual blaming practice when it comes to mental illness, especially in close relationships, which often admit far more complexity than such accounts acknowledge or allow (Pickard 2011; Brandenburg 2018; Svirsky 2020).

With regards to the normative challenge: some have raised concerns about the moral inadequacy of the Strawsonian recourse to 'exclusion' in cases of mental illness, especially insofar as withholding moral responsibility invariably involves withholding other forms of recognition and regard (Strawson [[1962] 1993 [1962] 1993; Shoemaker 2022; Glover 2014). David Shoemaker deems such exemption 'a form of morally objectionable discrimination against those viewed as having what I will label as *accountability disabilities*' (2022, 33). While Jonathan Glover refers to the intractability of exemption and exclusion as 'possibly insoluble', elaborating: 'blaming [people with relevant mental illnesses] may be unfair. But just as the reactive responses are at the core of relationships, exclusion from these responses is exclusion from part of that core' (2014, 274).

Our aim in the present paper is to respond to both of these challenges: to demonstrate (in response to the descriptive challenge) that Strawsonian frameworks can account for enormous complexity and nuance in interpersonal relationships affected by mental illness; and to show (in response to the normative challenge) that such frameworks are able to offer significant accommodation for mitigation on the basis of mental illness without recourse to asymmetrical relations.

In making this case, we will draw on the moral significance of difficulty in our assessments of responsibility, and particularly on what difficulty reveals about appraisals of adequate concern.[1] We will endeavour to sketch a theoretical terrain that reconciles what might superficially appear contradictory: that the basic demand can be both asserted, and in some respects met, even while someone (consciously and deliberately) fails to meet associated normative demands in their treatment of others. (We use 'normative demands' here to refer to demands of conduct. Superficially, the basic demand for due regard is closely related to normative demands regarding appropriate treatment. When X consciously and deliberately mistreats Y, we often assume this conscious and deliberate mistreatment *necessarily* reveals a lack of due regard. In the present paper, we will attempt to complicate this picture).

We understand this to be part of a broader recognition that while we are entitled to certain basic expectations within our moral relationships, it is also possible to expect *too much* of others (to have an inflated sense of what we are owed), and the demand that individuals meet certain normative demands with regards to their treatment of us – irrespective of the specific difficulties they face in meeting these demands – will at times be to demand more than we are owed.

Yet, as we will also acknowledge, there are limits to the significance of difficulty on Stawsonian accounts in reducing blameworthiness. We aim to defend the appropriateness of these limits. In §5, we distinguish between 'internal' and 'external' difficulty: where difficulty arises *from* lack of concern or motivation itself (i.e. internal difficulty) rather than from forces external to it (i.e. external difficulty) it does not have the same exculpatory strength. At such junctures the available accommodation diminishes, and we must either uphold blameworthiness, at least to some extent, or abandon equal and reciprocal relations. That is to say: exemption based on internal difficulty will necessarily have costs to symmetrical relations. These assessments are often complex, given the coexistence of internal and external difficulty, in various degrees, and we argue that this complexity tracks an appropriate ambivalence and multiplicity in our interpretation of these cases.

Importantly, we do not endorse any straightforward inference from the fact of mental illness to any particular assessment about moral responsibility (King and May 2018).[2] The framework we defend responds to certain dynamics, so in any given instance the crucial question will be 'what dynamics are present', rather than 'what condition is present'. Any particular assessment will therefore not only have to be considered on a condition-by-condition basis, but also on a case-by-case basis. That said, like many others within this debate, we are especially interested in how best to navigate complex cases, where agential capacity and incapacity coexist (and where it is not straightforward to say that someone was acting without knowledge or control in their harmful conduct, even if it is also true that their behaviour was affected by mental illness (Cf. Pickard 2017; Kennett 2003; Brandenburg 2018; 2019; Brandenburg and Strijbos 2020)). We take our

framework to be applicable to such complex cases, and will consider an example in the final section.

2. Excuse, exemption & agential status

An important distinction to consider within this debate is that between exemption and excuse. In general, excuse is taken to operate by indicating that what someone has done is not indicative of moral disrespect or insufficient regard (despite, perhaps, the way it may first appear). In this sense, it is compatible with the persistence of the basic demand, and excuse does not generally impute agential status in this deeper sense (though it can acknowledge other forms of incapacity). In contrast, exemption is an assessment that the basic demand *itself* is inappropriate, given the individual's lack of agential capacity. Exemption therefore inherently imputes agential status.

As others have pointed out (notably Brink and Nelkin 2013), it may be a mistake to over-invest in the distinction between excuse and exemption, or to treat them as wholly different classes. A complicating factor is that one can be exempt 'for a time'. That is to say: one can be temporarily ineligible for the demands of ordinary interpersonal relationships, but one can again be eligible for them at other times (or in other contexts). A transient incapacity may mean that the basic demand is temporarily out of place, but it should not deprive a person of membership in the moral community in an enduring sense. We can sometimes be temporarily ineligible for interpersonal demands that we are, at other times, able to meet. Here exemption encourages a temporary suspension of these demands, but not necessarily a sustained or permanent suspension. One can also be exempt with regard to a certain class of actions (i.e. locally exempt), but fully responsible with regard to others, depending on the specific agential incapacities on which the exemption is based. Given this potential variability in the duration and scope of exemption, there is perhaps more of a continuum between variations of excuse and variations of exemption than is often supposed when this distinction is evoked.[3]

Despite these qualifications, we will be drawing on this distinction going forward, and particularly on the relationship between exemption and the appropriateness of asserting the basic demand. Withholding blame on the basis of mental illness has often been viewed primarily through the paradigm of global or local exemption, where exemption requires withholding the basic demand, and therefore invariably implicates symmetrical relations (Shoemaker 2022). As noted above: to adopt the objective stance towards someone is inevitably to enter into a profoundly asymmetrical relationship with them. More recently, less extreme asymmetries have been proposed in the form of more accommodative stances, which might be compatible with the assertion of some forms of reproach or rebuke, even if they are not compatible with the robust forms of responsibility associated with fully symmetrical relations (Cf. Pickard 2011 (and many subsequent developments of 'responsibility without blame'); Brandenburg 2018; 2019; and Brandenburg and Strijbos 2020).

A popular comparison within this debate is that between individuals with mental illness and children.[4] The purpose of this comparison is to reveal that our interpersonal practice allows for a vast realm of moral engagement that does not rely on nor presuppose fully developed agency. But these comparisons also reveal, or concede, the

ongoing tension with regard to upholding symmetrical relations while making the sorts of allowances that the facts of mental illness seem to necessitate. There are many things we could say about the relationship between adults and children, which no doubt admit of great moral and interpersonal complexity despite our sometimes-reductive treatments of them, but it cannot be denied that these are not relationships of moral equality and reciprocity.

So even with regard to these more accommodative stances, our central dilemma asserts itself. If we err in the direction of attributing responsible agency, we will sometimes make the mistake of blaming over-readily. Many people are so averse to this prospect that they prefer the alternative. But if we err in the direction of withholding responsibility, we will sometimes make the mistake of needlessly undermining a symmetrical relationship. The normative challenge, which we described in §1, emphasised that this sort of withholding is not something we should resort to without conflict as the lesser of evils. There is a great deal at stake in such assessments. As Smith writes:

> For anyone who has had the unpleasant experience of having her emotions or reactions dismissed in this way, it should be clear that being denied responsibility for one's attitudes has its costs. Such denials can be deeply patronizing and disrespectful, and we should not be too eager to resort to them, either in our own case or in our treatment of others. (2005, 269)

To answer the normative challenge it is therefore necessary to find different ways to describe the bases of diminished responsibility that arise in response to mental illness. As we will argue in what follows: an appreciation of the moral significance difficulty on Strawsonian frameworks, and the relevance of difficulty to assessments of adequate concern, enables us to navigate some of the dilemmas, contradictions and complexities generated by interpersonal relationships affected by mental illness. In the next section, we will explore these dynamics as they pertain to 'external difficulty', or difficulty that arises from outside of moral concern itself. We suggest that the accommodation that is possible on the grounds of external difficulty, while simultaneously maintaining symmetrical relations, is considerable, including with regards to mental illness. We explore this scope further in §IV, when we consider the role (and the possibilities) of interpretative generosity.

In §5, however, we turn to the limits of this accommodation. Here we turn to 'internal difficulty', which arises from insufficient concern itself. As we argue: unlike external difficulty, exculpation on the basis of internal difficulty *necessitates* withholding the basic demand, and therefore necessitates forms of relational asymmetry and moral exclusion. We argue, however, that at this point the asymmetries are justified, and the normative challenge is adequately answered.

3. The significance of difficulty

Difficulty is an important consideration in assessments of moral responsibility. When someone fails to do the right thing, but it would have been incredibly difficult for them to do it, we often think that the difficulty they faced is blame-mitigating. But while exempting on the basis of incapacity is quite straightforward (you are not responsible, because you were unable to comply), the significance of difficulty is more complicated: after all, when something is difficult, it is not impossible. Through what means does

difficulty diminish responsibility? In what respect does a wrong action (perhaps deliberately and consciously undertaken) become less culpable merely because it was difficult to avoid?[5]

One powerful explanation relies on Strawsonian frameworks (Hartford, 2022). Difficulty allows us to see that wrongful conduct is not necessarily indicative of inadequate concern. When it has this implication, it is a version of what we are calling 'external difficulty', which will be our focus until §5.

Let us begin with a simple example in order to describe the dynamics involved. Let's say Das is deeply afraid of the dark, but her little brother is severely ill and her mother has requested that she cross a dark field at night to urgently retrieve some medication. Here it might simultaneously be true that Das is capable of such a journey (there is no question that she would run through the field if chased by an assailant, for instance, or while fleeing a fire) while it is nevertheless true that undertaking it will be exceptionally difficult for her depending on how excessive her fear is. Should Das set out on this mission and fail, her blameworthiness for this failure would seem significantly mitigated on account of the unique difficulty she faced given her fears. Furthermore, and crucially for our purposes, the sense in which her failure implied a lack of concern for her little brother and her mother would also be significantly reduced. Depending on how overwhelming her fear was, it could simultaneously be true that the interests of her mother and her little brother mattered deeply to her, and yet that she nevertheless failed to retrieve the medication even though she was capable of doing so.

The crucial dynamic here is that the basic demand can be asserted, and indeed *met* (Das can possess appropriate regard for those involved) while she nevertheless fails to meet the associated normative demands on account of the unique difficulties she faced. Furthermore, this can remain true *even if* her failure is ultimately deliberate, controlled and consciously chosen; even if, for instance, she made a firm decision, at some point, not to persevere further, but rather to relent and give into her fear. In this respect her failure could arise *from her agency*, to a significant degree, but the conditions and circumstances under which she exercised this agency could nevertheless be substantially mitigating, insofar as they revealed that her failure to act appropriately was not necessarily indicative of a failure to care appropriately.

Insofar as the basic demand is the bedrock of symmetrical relationships, acknowledgement of the moral significance of difficulty therefore allows that in many cases the basic demand can be asserted, and symmetrical relationships sustained, even where someone fails to meet associated normative demands in their conduct (and even, in some cases, where they do so deliberately and with control). Das is not rendered outside of the moral community on account of her fear; her mother and brother need not adopt a different sort of interpersonal stance towards her. Rather, the facts about her experience become important in analysing the content of her attitudes, and ultimately allows them to diminish affective blame without diminishing symmetrical relations.

Part of what is acknowledged here, in the recognition of Das's situation, is that while we are each entitled to robust demands from other members of the moral community, it is also possible to demand *unfairly*, or to be overdemanding in what we ask of others. Hanna Pickard talks about 'entitlement' as a foundational aspect of evaluations of moral responsibility and blameworthiness. And in our discourse about what we

'demand' in our interpersonal moral relationships, we can see the resonance of this entitlement. Yet what we have also been hoping to explore (and will elaborate on further in the next section) is that one can be entitled – one can assert the basic demand – without being *over entitled*, or without demanding in a way that is insensitive to and unresponsive to the challenges and difficulties another agent encounters.

4. Responding to the normative challenge: the role of interpretative generosity

If responsibility is ultimately an assessment of another agent's pattern of moral concern, the endeavour to establish an individual's responsibility necessitates the endeavour to occupy their standpoint and to understand how their subjective experience impacts on our assessment of what constitutes an objectionable attitude, and (in turn) of whether such attitudes were genuinely evinced in their conduct. This can never be perfectly achieved; we can never genuinely know someone else's experience. Nevertheless this practice is indispensable to all forms of interpersonal evaluation. It is the 'interpretative project', through which we ascribe intentions, desires, wishes, attitudes, feelings and beliefs. Within this project, all agents wish to be appraised with generosity, especially with regards to their moral failures (as when we desperately implore others, one way or another, to understand where we were coming from in our harmful or wrongful conduct).

This practice potentially has vast application to cases of mental illness. Moreover: in conjunction with an appreciation of the significance of difficulty for assessments of sufficient concern, it may allow us to recognise that the deepest, most fundamental dynamic of symmetrical relations is in fact compatible even with profound failures to meet normative demands on the basis of mental illness.

Our emphasis on interpretative generosity in such relationships – and the moral importance in attempting to undertake the interpretative project even in our relationships with people with profoundly different experiences from ourselves – aligns with other work, such as Jonathan Glover's and Jeannette Kennett's, which emphasises the importance of the project of interpretation when it comes to disordered minds.

It is easy enough to imagine fear of the dark (we have probably all experienced a variety of it); it is much harder to imagine, let alone inhabit, the sorts of alien terrors, pressures, beliefs, sensations and perceptions that might arise in disordered minds. As Glover's project indicates, it is no simple matter to put ourselves in the shoes of people with severe psychiatric problems. In one respect, it seems presumptuous to even suggest it. And yet the refusal to make the attempt could be all the more harmful: deciding that someone's experience is unreachable and uninterpretable is a profound form of disengagement with that person; to opt to turn away from 'the gulf that defies description', rather than endeavouring to cross it, or rather than being open to believe that even the most seemly inexplicable behaviour might actually be explicable if it were understood properly within its circumstances.[6] As Kennett writes, on Glover: 'we extend [our normal framework of interpretation] by being open to the thought that others' reasons might be very different given their different circumstances – and features of a disorder might provide very different circumstances' (2017, 741).

The effort to understand the subjective contexts that other people inhabit is indispensable to the project of interpersonal interpretation generally. The more different those

subjective contexts are, the more important it is to attend to the testimony and experience of people affected by various disorders, and to endeavour to cultivate the understanding we lack (Kennett 2017). This is not a simple undertaking. As Kennett notes, the project of interpretation as applied to disordered minds requires both moral wisdom and commitment, as well as 'humility, charity, open-mindedness, patience, carefulness, respect and good will' (2017, 742).

Returning to the significance of difficulty: better recognition and understanding of the phenomenology of various conditions, and the sorts of obstacles a person potentially faces in meeting normative demands, can therefore help to preserve and extend, rather than undermine, symmetrical relationships.

Insofar as we succeed in extending the project of interpretation to disordered minds, we might similarly discover that in certain circumstances – contrary to first appearances, perhaps – it is possible that someone had sufficient concern *even if,* in the scheme of what they were called upon to manage and endure, that concern was unable to manifest in their conduct. We need not *withdraw* the basic demand, in such circumstances, in order to consider them less blameworthy.

With this in mind, let us return to the normative challenge we introduced in §1, which charged Strawsonian accounts with over-readily resorting to asymmetrical relations in order to navigate cases of mental illness. As we have argued: the accommodation that is available to withhold blameworthiness on Strawsonian accounts, while simultaneously upholding symmetrical relations, is in some respects as vast as we are able to make it. The more we are able to practice the virtues that Kennett describes, the greater the possibility of retaining symmetry even in relationships that are severely affected by mental illness. Many of us will fall far short of what is possible here. But if we fall short, it is not an indication that such accommodations were not *possible* within the more nuanced Strawsonian framework we are proposing. They are possible, and the framework does not *necessitate* any quick recourse to exclusion, or to exemptive stances.[7]

There are various strengths to this approach. Firstly, it is able to recognise, and take seriously, the burdens and difficulties an individual might face on account of mental illness – and to explain the basis of diminished or withheld blame on these grounds – without necessarily resorting to asymmetrical relations. Secondly, it is co-extensive with our normal practice of interpersonal interpretation, and does not require the adoption of a different stance, or a different practice, in order to make the necessary accommodation.

Thirdly, this approach is particularly applicable to complex cases where agential capacity and incapacity coexist. Being 'incapable' can be interpreted as a threshold concept: either you are beneath the relevant threshold, or you aren't. But there are many cases where it is inaccurate to characterise particular conditions as generating 'incapacity' in the binary sense. In many cases there will be much more knowledge, deliberation and control than such wholescale incapacitation allows. Yet it is not only at extremes that the grounds for diminished blame become compelling, and we can draw on scalar notions of the relevant capacities as affected by different impairments. As Brink and Nelkin write: 'insanity and immaturity are clearly scalar concepts. But so are the ideas of paralyzing emotions, irresistible urges, and a disabling depression. So too is the hardness of choice' (2013, 309). Emphasising the significance of difficulty, and its relevance to

appraisals of sufficient concern, allows a scalar approach that is responsive to these more ambiguous cases.

Relatedly, as Sofia Jeppsson has pointed out, paradigms of exemption (and associated asymmetrical stances) are often hard to reconcile with self-appraisal from the perspective of the affected individual (Jeppsson 2021; 2022). It is awkward to occupy a state of 'self-exemption' on the basis of mental illness, especially for decisions that you recognise that you consciously participated in and experienced as involving a degree of deliberation and choice at the time. Under such circumstances, it is hard to think of *oneself* as a non-agent, or outside of the moral community, in the relevant sense. Yet, particularly when agency is exercised under immense strain, pressure or difficulty, it is nevertheless possible to extend self-compassion and reduce self-blame.

For all we have said, however, it remains the case that there comes a point where this difficulty-oriented approach will be insufficient, and where a different (non-participant) stance becomes necessary (or, alternatively, an assertion of blameworthiness). Our present objective has been to establish that this is not a near point, nor one we need to hasten to in order to avoid being overly entitled in our interpersonal moral relationships affected by mental illness. But it is also interesting to consider the limits of this accommodation on Strawsonian grounds, and where they leave us; we turn to this in next section.

5. Varieties of difficulty & the limits of accommodation

As we have explored in the previous two sections, the fact that someone faces particular difficulty (and potentially extraordinary difficulty) in meeting particular normative demands seems to be of immense relevance in our appraisals of their culpability for the relevant failures. But, importantly, the fact of difficulty does not universally reduce blameworthiness.

Return to frightened Das: the difficulty she faces is *external* to her moral concern. For this reason, it allows us to recognise that her failure to act as she ought does not necessarily imply a failure to care as she ought. Matters become more complicated when we consider difficulty that is not external to an agent's moral concern, but rather *internal* to it.

Again, let's look at a simple example in order to describe the dynamics involved. Imagine that Das has another brother, Zed. Zed is not afraid of the dark, but when he is asked to help, he too does not comply. Seeing his mother's worry, his brother's illness, he feels nothing but irritation (his brother, as far as he's concerned, is just trying to be the centre of attention). Now Zed also faces difficulty here: it's very difficult to do what you feel no reason, impetus or desire to do. It's difficult to act as though you care when you don't care. However, unlike the case of Das, this difficulty is *internal* to Zed's moral concerns (it arises from his lack of moral concern itself) rather than from factors external to it. This fact changes matters fundamentally: and while most of us would feel that the difficulty Das faced mitigates her blameworthiness, few of us would feel (all else remaining equal) that the difficulty Zed faced would mitigate his.

While *external* difficulty complicates the inference from moral conduct to moral concern, *internal* difficulty does not. How to explain this duality? Again Strawsonian accounts are particularly explanatory: since blame is justified in response to insufficient

concern, difficulty does not mitigate when it stems from (or exacerbates) lack of concern *itself*.

Let's further imagine that the difficulty Zed faces on this occasion is attributable, in part, to an impairment of sorts. Perhaps he has a combination of personality disorders that contribute to his lack of concern in this case. If we isolate this aspect, and stipulate that it is on the grounds of such an impairment that Zed fails to do what he ought, what position are we placed in with regard to his blameworthiness for this failure? Zed has *not* met the basic demand, and there is no sense in which the difficulty he faces indicates that, despite appearances, he has.

Some philosophers argue that such internal difficulty should also be relevant to mitigation, and have argued on this basis against Strawsonian accounts (Rosen 2008; Lamb 2022). After all: if we bear in mind that the relevant impairments are beyond Zed's control, and that they generate significant difficulty for him in meeting normative demands, surely we should be as understanding about his circumstances as we are about Das's?

For our present purposes, we needn't take a side in this particular dispute. Our point is not that someone who faces internal difficulty is *always* blameworthy, no matter the reason. Rather, and crucially, our point is that while withholding blame on account of *external* difficulty is compatible with symmetrical relations (as we have argued at length above), withholding blame on account of *internal* difficulty necessitates the adoption of asymmetrical relations. In determining that the basic demand *cannot fully be made* of Zed, we therefore abandon the central tenet of equal and reciprocal moral relations. Where blame is withheld on the basis of internal difficulty it therefore must be, at least to an extent, at the expense of symmetrical relations.

Earlier we described interpretive generosity as a key feature of our interpersonal assessments: we can consider the sorts of obstacles an agent faced in their failure to meet normative requirements; we can try (in our imperfect way) to see how things were from their perspective, and often this process will lead us to reconsider our initial appraisals. This undertaking would be effective in the case of Das, but it would have a very different implication in the case of Zed. If we tried to understand where Zed was coming from, and correctly perceived his profound indifference, we would hardly, on account of this perception, come to feel that he was not a fault for his failure to assist.

We do not mean to imply that diminished blame and asymmetrical relations are the only appropriate ways to respond to Zed. There is also a case to be made for the full assertion of blameworthiness. It is important to acknowledge and affirm the humanity of people on all sides of these relationships. In §3, we warned against over-entitlement in what we demand of others, but it is also possible to recognise the point at which entitlement is legitimate, and with it: blame, anger and disappointment. There is a point at which it is too much to be required to 'rise above' ill treatment that is founded in ill regard, and there is moral value to asserting that a person in this position is entitled to feel hurt, wounded and angry. In their best incarnations, the emotions of blame and moral anger can affirm self-respect and self-protection, and condemn contempt and harm.

Furthermore, given the interpersonal costs of asymmetrical relations, there may also be moral defences for the assertion of blameworthiness in such cases. This is especially so if we view responsibility and blame as intractably bound to the web of other reactive

attitudes that constitute the core of our relationships (Glover 2014). If such intractability holds, the person from whom we withhold blame (despite their genuine moral disregard for us) is also the person from whom we ultimately withhold much else; we do not, in Christine Korsgaard's phrase, 'risk [our] happiness or success' on those who we cannot ask to care for our humanity (1992, 306). While all sorts of relationships may be possible with this person, these relationships will invariably be shallower than those which are bound to mutuality, reciprocity and answerability. Asserting blame, and asserting this mutuality, may therefore also be a bid for this deeper sort of relationship, rather than just being a cruel and pointless expression of hostile emotion. This is particularly true when these dynamics emerge in our closest personal relationships: between partners, siblings, children and their parents, or old friends.

On the framework we are advancing, the extent of this asymmetry can admit of degrees. This is distinct from conceptions which treat exemption as an all-or-nothing affair.[8] This scalarity is also a response to the normative challenge: in addition to the expansive accommodation for sustaining symmetrical relations despite withholding blame on account of mental illness that we argued for in §3 & §4, we can also argue for the possibility of degrees of asymmetry even at the point where this accommodation ceases, and where withholding blame necessitates asymmetrical relations.

We acknowledge, however, that even these lesser asymmetries carry interpersonal costs. We further acknowledge that there is a point – pertaining specifically to internal difficulty – where a Strawsonian account demands that we either assert blame, or else adopt asymmetrical relations. Our goal over the preceding three sections– in answering the normative challenge– has been to demonstrate that this is by no means a near point, and that there is substantive room to manoeuvre (and to both withhold blame and uphold symmetrical relations) before it is reached.

In the next and final section, we will turn to the descriptive challenge: the sense in which such an account could align with our actual blaming practice in the context of close relationships.

6. Responding to the descriptive challenge

To close, it will be helpful to consider the proceeding discussion in the context of a specific case. As we said at the outset: the framework we defend responds to certain dynamics rather than to any specific condition. In the case of any particular condition or combination of conditions, different confluences of the relevant dynamics could be at play:the sense of contrast and continuity between self and illness might be different in each case, as might the mix of internal and external difficulty, and their inculpating and exculpating forces.

Let us consider a marriage affected by the husband's substance addiction. Let's say the husband has, over the years, had various lapses into severe substance abuse, as well as periods of recovery. He has also been, at times, selfish and self-interested in ways that aren't straightforwardly attributable to any disorder. Let's suppose that during a relapse, the husband accrues significant financial debt that both he and his wife are legally liable for. Let us grant that in accruing this debt in her name, he has wronged her.

One way for the wife to respond to this state of affairs is with resentment and anger; that would be the natural reaction in a symmetrical relationship. However, we might think

that she ought to temper these affective reactions given the facts of his substance addiction, and the sense in which it impaired his control over his conduct. From one vantage, if we emphasise the sense in which addiction is a disease *et cetera*, a Strawsonian perspective might recommend that the wife suspend the basic demand and associated forms of responsibility for her husband's harmful behaviour, and detach from affective blame responses to the harm that he has caused her.

The normative and descriptive challenges we described in §1 would potentially apply to such a 'straightforward' application. The normative challenge would emphasise the costs of the switch to asymmetrical relations, especially in such an intimate context. The point of this challenge is not necessarily to assert that all asymmetries are wrong: asymmetrical relationships take many forms, and in some cases, they will be entirely appropriate. However, there are also costs of these asymmetries, particularly in certain contexts. Having the same stance to your husband (for instance) as you do towards an irresponsible teenage son is not without interpersonal loss. Something might be gained in adopting such a stance (perhaps it allows you to better manage affective blame), but something valuable is also lost.

On the other hand, the descriptive challenge would counter that such an application simply does not capture the complex interpersonal dynamics at play in such a case. In practice, there is often no easy recourse to 'detachment' to manage affective blame when you've been wronged by someone very close to you, and this is especially so when their conduct, even if impacted by disorder, was also knowing and deliberate in other respects.

On the more nuanced application of the Strawsonian framework we are proposing, to the extent that the husband's addiction generated *external difficulty* relevant to his wrongdoing – i.e. external to his moral concern for his wife – it is *possible* for her to mitigate or exculpate without abandoning symmetrical relations. Insofar as his addiction generated external difficulties (through intense cravings, withdrawal symptoms, other overwhelming stresses and pressures *et cetera*), and insofar as his wife was able to access the necessary 'charity, open-mindedness, and patience' in interpreting the impact of these external difficulties on his wrongful conduct, there is potentially enormous scope for her to mitigate blame while still upholding symmetrical relations. I.e. There is scope for her to recognise that his conduct, while harmful to her, was not incompatible with her assertion of the basic demand within their relationship. In response to the normative challenge, therefore, there need not be any easy or quick recourse to asymmetrical relations in order to make the relevant accommodations.

However, insofar as his addiction also generated *internal difficulty* – i.e. insofar as it genuinely made him less concerned for her, and more inclined to disregard her rightful claims and interests – it is much harder for her to mitigate on account of his addiction while retaining symmetrical relations. Part of the way in which his addiction has manifested is precisely in his lack of concern for her, and in withholding blame despite this, she invariably has to cease asserting the basic demand, at least to a certain extent.

In many cases all of these forces will exist at once. Part of our aim is to draw out the immense complexity and ambiguity in close relationships affected by particular disorders, but also affected by all the difficult dynamics that emerge in any close relationship. It is not always perfectly apparent where the features of a disorder begin and where other more enduring parts of a person's nature end. It is not always clear when someone

was acting under immense external difficulties, or when their conduct actually evinced disregard and disdain for those they hurt, or in which combination these different forces existed.

This multiplicity can generate deeply complicated, and even ambivalent, assessments when it comes to appraising the conduct of particular individuals. While the forces we have been describing are relatively simple when abstracted, their manifestation in reality is anything but. Intractable epistemic difficulties remain: how can we really know what obstacles someone faced in meeting normative demands? How can we know the source of those obstacles, and their nature? How can we appraise whether they were external or internal to moral concern and motivation itself?

Sometimes these determinations will be relatively straightforward, but often they will be very hard, and it will be very hard to determine how to weigh these various factors. Even in our relationships with a specific individual, it's something we might come back to over and over again (as we interpret, and reinterpret, the different forces at play; or as future conduct leads us to reappraise the past). The complexity of these judgments is a far cry from the detachment of the objective stance; it demonstrates the descriptive nuance that is available on Strawsonian frameworks, and the many alternatives to simple detachment on such paradigms. This reflects the complexity often evinced within inter-personal relationships affected by mental illness, where people grapple with the contra-dictory spaces that both recognise and acknowledge our interpersonal entitlements, but also recognise and acknowledge the limits of these entitlements.

Notes

1. We focus here on agent-relative notions of difficulty, where something is difficult in the rel-evant sense insofar as it is difficult *for* a particular agent (Cf. Bradford 2017 and von Kriegstein 2019 for further debate on the agent-relativity or agent-neutrality of difficulty).
2. King and May (2018) have even argued that mental illness can sometimes enhance responsibility.
3. 'Scope' and 'duration' are taken from Brink and Nelkin (2013); they propose that exemptions should be considered 'as comparatively global or standing excuses.'
4. Svirsky looks at both an addict in one case, and a child in another, to show how our personal relationships with so-called 'marginal agents' impacts blameworthiness and standing to blame (Svirsky 2020). Pickard compares the detached clinical stance she encourages to 'good parenting,' and acknowledges that there is an 'inherent power imbalance' between clinicians and service users, as there is between children and parents. Brandenburg also draws on the comparison in developing her nurturing stance and indicates that 'carers can be seen to respond to children as young as two or three in a nurturing manner'; elsewhere she makes a sustained comparison between our relationship with children and the relation-ship of clinicians and service users (Brandenburg 2019).
5. These and related questions about the significance of difficulty in agential appraisals have been the subject of growing philosophical attention (Cf. Nelkin 2016; Bradford 2017; von Kriegstein 2019; Hartford 2022).
6. To use Glover's quotation of Karl Jaspers (Glover 2014).
7. Work on moral responsibility navigates normative and naturalistic realms. In responding to the normative challenge, we are drawing on the normative possibilities of the account we are supporting, rather than only on our blaming practice.
8. Our account therefore aligns – in these cases of exemption on the basis of internal difficulty – with theories which have proposed more accommodative exemptive stances, and do not

view the basic demand as an all-or-nothing affair (Cf. Brandenburg 2018). We hold, however, that there remain interpersonal costs even to less extreme asymmetrical stances.

Acknowledgements

Disclosure statement

Funding

Notes on contributors

Anna Hartford is a writer and researcher based in South Africa. She is currently a postdoctoral fellow at the Neuroscience Institute at the University of Cape Town.

Dan J. Stein is Professor and Chair of the Department of Psychiatry and Mental Health at the University of Cape Town, and Director of the South African Medical Research Council's Unit on Risk & Resilience in Mental Disorders. His most recent book is *Problems of Living: Perspectives from Philosophy, Psychiatry, and Cognitive–Affective Science.*

References

Bradford, Gwen. 2017. "Hard to Know." In *in Responsibility: The Epistemic Condition*, edited by P. Robichaud and J. W. Wieland, 180–198. Oxford: OUP.

Brandenburg, Daphne. 2018. "The Nurturing Stance: Making Sense of Responsibility Without Blame." *Pacific Philosophical Quarterly* 99 (S1): 5–22. https://doi.org/10.1111/papq.12210

Brandenburg, Daphne. 2019. "Inadequate Agency and Appropriate Anger." *Ethical Theory and Moral Practice* 22 (1): 169–185. https://doi.org/10.1007/s10677-019-09982-w

Brandenburg, Daphne, and Derek. Strijbos. 2020. "The Clinical Stance & the Nurturing Stance: Therapeutic Responses to Harmful Conduct by Service Users in Mental Healthcare." *Philosophy, Psychiatry & Psychology* 27 (4): 379–394. https://doi.org/10.1353/ppp.2020.0049

Brink, David O., and Dana K. Nelkin. 2013. "Fairness and the Architecture of Responsibility." In *in Oxford Studies in Agency and Responsibility*. Volume 1, edited by David Shoemaker, 284–314. Oxford: Oxford University Press.

Glover, Jonathan. 2014. *Alien Landscapes? Interpreting Disordered Minds*. Cambridge, MA: Harvard University Press.

Hartford, Anna. 2022. "Difficulty & Quality of Will: Implications for Moral Ignorance." *Philosophical Explorations* 25 (2): 141–158. https://doi.org/10.1080/13869795.2022.2042585

Jeppsson, Sofia. 2021. "Exemption, Self-Exemption and Compassionate Self-Excuse." Presentation at the "Responsibility, Psychopathology and Stigma" Conference, NeuroEpigenEthics, Antwerp.

Jeppsson, Sofia. 2022. "Solving the Self-Illness Ambiguity: The Case for Construction Over Discovery." *Philosophical Explorations* 25 (3): 294–313. https://doi.org/10.1080/13869795.2022.2051589

Kennett, Jeanette. 2003. *Agency and Responsibility: A Common-Sense Moral Psychology*. Oxford: Clarendon.

Kennett, Jeanette. 2017. "Morality and Interpretation: Commentary on Jonathan Glover's *Alien Landscapes?*" *Journal of Applied Philosophy* 34 (5): 737–742. https://doi.org/10.1111/japp.12209

King, Matt., and Joshua. May. 2018. "Moral Responsibility and Mental Illness: A Call for Nuance." *Neuroethics* 11 (1): 11–22. https://doi.org/10.1007/s12152-017-9345-4

Korsgaard, Christine M. 1992. "Creating the Kingdom of Ends: Reciprocity and Responsibility in Personal Relations." *Philosophical Perspectives* 6: 305–332. https://doi.org/10.2307/2214250

Lamb, Matthew. 2022. "Quality of Will Accounts and Non-Culpable Mental Disorders." *Journal of Ethics and Social Philosophy* 22 (3): 415–421.

Nelkin, Dana K. 2016. "Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness." *Nous (detroit, Mich)* 50 (2): 356–378. https://doi.org/10.1111/nous.12079

Pickard, Hanna. 2011. "Responsibility Without Blame: Empathy and the Effective Treatment of Personality Disorder." *Philosophy, Psychiatry, Psychology* 18 (3): 209–224. https://doi.org/10.1353/ppp.2011.0032

Pickard, Hanna. 2017. "Responsibility Without Blame for Addiction." *Neuroethics* 10 (1): 169–180.

Rosen, Gideon. 2008. "Kleinbart the Oblivious and Other Tales of Ignorance and Responsibility." *Journal of Philosophy* 105 (10): 591–610. https://doi.org/10.5840/jphil20081051023

Scanlon, T M. 2008. *Moral dimensions: Permissibility, meaning, blame*. Harvard University Press.

Shoemaker, David. 2022. "Disordered, Disabled, Disregarded, Dismissed: The Moral Costs of Exemptions from Accountability." In *Agency in Mental Disorder: Philosophical Dimensions*, edited by M. King and J. May, 33–62. Oxford: Oxford University Press.

Smith, Angela M. 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115 (2): 236–271. https://doi.org/10.1086/426957

Strawson, Peter F. [1962] 1993. "Freedom and Resentment." In *Proceedings of the British Academy*, edited by J. M. Fischer, and M. Ravizza, 48, Reprinted in 1–25. Ithaca, New York: Cornell University Press.

Svirsky, Larisa. 2020. "Responsibility and the Problem of So-Called Marginal Agents." *Journal of the American Philosophical Association* 6 (2): 246–293. https://doi.org/10.1017/apa.2019.36

von Kriegstein, Hasko. 2019. "On Being Difficult: Towards an Account of the Nature of Difficulty." *Philosophical Studies* 176 (1): 45–64. https://doi.org/10.1007/s11098-017-1005-3

Watson, Gary. 1993. "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme." In *Perspectives on Moral Responsibility*, edited by J. M. Fischer and M. Ravizza, 119–148. Ithaca, New York: Cornell University Press.