The *BAD* Theory of Blame

What does blame add to belief? Some argue that blame is a set of connected beliefs, reproach or sanction, or a set of reactive emotions. George Sher developed an original and compelling theory of blame as a belief/desire pair. I argue that Sher's account must be revised to better align with common experience. His account, while elegant and parsimonious, cannot coherently understand exemptions and is too readily compatible with the objective attitude. I suggest adding a missing element that makes sense of exemptions, renders blame incompatible with the objective attitude, and aligns with common experience—the *BAD* theory of blame.

Keywords: Blame, George Sher, Belief, Desire, Affect, Excuse, Exemption, Objective Attitude

---

> ***Bargle***: I see that I can never hope to refute you, since I no sooner reduce your position to absurdity than you embrace the absurdity.
>
> ***Argle***: Not absurdity; disagreement with common opinion.
>
> ***Bargle***: Very well. But I, for one, have more trust in common opinions than I do in any philosophical reasoning whatever. In so far as you disagree with them, you must pay a great price in the plausibility of your theories.
>
> ***Argle***: Agreed. We have been measuring that price. I have shown that it is not so great as you thought; I am prepared to pay it. My theories can earn credence by their clarity and economy; and if they disagree a little with common opinion, then common opinion may be corrected even by a philosopher.

Bargle: The price is still too high.

Argle: We agree in principle; we're only haggling.[1]

What does blame add to belief?[2] When we blame someone, what does our blame contain beyond a mere belief that the person acted wrongly? Some philosophers argue that blame is merely a set of connected beliefs, simply reproach or sanction, or essentially a set of reactive emotions. George Sher has developed an original and compelling theory of blame, arguing that blame is best understood as an interconnected belief/desire pair.

In this article, I will adopt the role of Bargle, arguing that Sher's prominent theory of blame, while exquisitely elegant and parsimonious, must be revised to better align with common experience. I believe Sher's account is missing a crucial element that, when added, secures a fully satisfying theory of blame. I want to be clear at the outset, though, that this critique truly is a friendly one—I do not wish to reject or replace Sher's account of blame, but to constructively add to it; the *BAD* theory of blame will take its cues from Sher's original account. I should also clarify that this essay aims to investigate the *nature* of blame and avoids questions of blameworthiness, standing to blame, and blame's normative place within our lives. While those questions are well-worth investigating, here, I will attempt to understand what blame is and how it works. I follow Sher's lead in directing our inquiry towards blame as a common "phenomenon in the world. We

[1] Lewis & Lewis, "Holes," 5-6.
[2] Sher, *In Praise of Blame*, 75.

have all had the experience of blaming others and being blamed by them, and what we are trying to find out is what such blame amounts to."[3] With this orientation in mind, let us begin.

Sher's 2006 book, *In Praise of Blame*, was one of the first detailed treatments of blame as a robust phenomenon. Chapter six of the book provides a comprehensive look at Sher's theory of blame. Below, I will lay out the essential features of the view.

According to Sher, blame consists of a belief/desire pair. One must believe that another has either acted wrongly or demonstrates a vicious character—both are presumably blameworthy transgressions.[4] Moving forward, B will represent the belief that either someone acted wrongly or has a vicious character. Paired with this belief is a desire that the person "*not have done* what he in fact did" or "that the person's character not be as bad as it in fact is."[5] The first desire is a backward-looking one; the second looks at the present. Either desire will be represented with D.

B and D are connected in that they both focus on a normative[6] transgression, and D's presence is explained by B's content. It would make little sense for me to desire that you had not performed an action, *because* it was wrong, unless I believed the act was indeed wrong. Thus, blame differentiates itself from mere frustration of one's non-normative preferences.

This B/D pair creates a *disposition* towards experiencing negative emotions typically associated with blame. Importantly, though, B and D only need create a *disposition* towards anger, resentment, indignation, and the like. For a complete instance of blame, on Sher's account, B and D need not manifest any actual experience of negative emotion.

---

[3] Sher, *In Praise of Blame*, 112.
[4] Sher, *In Praise of Blame*, 95.
[5] Sher, *In Praise of Blame*, 102.
[6] Sher specifically states that blame responds to *moral* transgressions and takes Chapter 7 to spell out blame's necessary connection to morality specifically. For the purposes of this paper, though, I will leave open the possibility of being motivated to blame in any normative domain; epistemic, social, practical/prudential, professional, etc.

Every instance of blame involves a B/D pair. That pair often creates a disposition towards experiencing negative emotions.[7] Perhaps for creatures like us, though, having a genuine B and D always creates a *disposition* towards anger. That disposition often manifests in actual negative emotions—perhaps most of the time. And perhaps a smaller subset of those emotions creates expressions of blame—yelling, reproaching, withdrawing, etc. The disposition is jointly explained by both B and D. B provides the disposition its content—we blame others *for* specific transgressions—and D motivates the disposition—we are often prone to anger when we blame. Often, B and D are working below the level of conscious awareness; I might spring into a 'blaming state' without explicitly determining B and then instantiating D. We probably rarely hold these thoughts explicitly in our inner monologue.

I find Sher's account of blame genuinely compelling. It is elegant and concise, clearly superior to rival accounts, and explains most common experiences regarding blame. However, I believe that Sher's account requires a bit of fine-tuning before it can be readily sold in local supermarkets; I believe his account lacks a crucial element, preventing it from capturing an array of common experiences. In this paper, I aim to illuminate this issue and provide its remedy. I will propose a tradeoff for Sher's account of blame: I will add an element to blame, slightly reducing its elegance and parsimony, and in exchange, I can offer a better fit with common experience.

Here is how my proposed trade will proceed: in section 1, I will review blame's common experiential explanandums and theoretical desiderata and explain how Sher's account secures these requirements. Where relevant, I contrast Sher's theory with its rivals and highlight the clear advantages of Sher's account. This is relevant to my project because I do not wish to reject Sher's

[7] Sher, *In Praise of Blame*, 96, 102, 103, 104-105, 115-116.

view, merely to reform it. In section 2, I introduce and explain two forgotten blame explanandums and one blame desideratum. I argue that these additions are critical to developing a complete theory of blame. In section 3, I argue that Sher's theory struggles to make sense of one new blame explanandum and one new blame desideratum. This creates a problem for Sher's view because any satisfying theory of blame should account for these additions. In section 4, I propose the *BAD* theory of blame and argue that this revised version of Sher's theory can make sense of our two additions while better aligning with common experience. Section 5 addresses objections.

Section 1: Theory Construction

What should a theory of blame *do*? What theoretical values and common experiences should an account of blame capture? George Sher and R. Jay Wallace have conveniently compiled various requirements for any proposed theory of blame. We can split these requirements into two general categories: *blame explanandums* and *blame desiderata*.

Blame explanandums are phenomena that we commonly associate with blame. Where blame shows up, these phenomena tend to follow. Blame explanandums are Bargle's talking point; because of their close association with blame itself, blame explanandums are experiences that must be *explained* by a successful theory of blame. Blame desiderata align with Argle's emphasis on theoretical values: they are analytic, philosophical requirements of any theory—clarity, parsimony, elegance, etc. Blame desiderata are partially motivated by common experience, but they are centrally a philosopher's focus. Both are critical to developing a successful account of blame.

Below, I will briefly introduce and explain four blame explanandums and five blame desiderata, lifted directly from the literature on blame-theorizing, and explain why Sher's account

can easily make sense of each. I will also contrast Sher's B/D pair with its rivals, highlighting the advantages of Sher's account in order to explain my focus on *his* theory over its competitors.

Blame Explanandums—Provided by Sher (2006)[8]

1. *Anger, etc*. Anger, indignation, bitterness, contempt, resentment, rage, etc. are all typical emotions commonly experienced when we blame others; we often direct these emotions towards our malefactors. While Sher maintains that anger and its relatives are not essential to blame, they are commonly experienced as a result of the B/D pair. On Sher's account, blame's common angry emotions are explained by the content of B. My belief that you did something wrong explains why I might become resentful towards *you* for *that* transgression.

On an expressivist Strawsonian account, however, blame *just is* one's directed negative emotions.[9] To blame someone is simply to experience anger, etc. towards him. But the Strawsonian account is unable to explain why negative emotions often result from being wronged. Sher's answer is that they depend upon a *belief* that one was wronged. Thus, changes in B can dissipate negative emotion; if I come to discover that you had a legitimate excuse for your actions—perhaps you tripped and stepped on my hand—or that *you* did not perform the wrong, my anger will dissipate while my pain remains. But the relevant explanation of why this occurs is that I have come to *believe* that you did not act wrongly.[10]

Even adding B to negative emotion, as Wallace's Reactive Account does, leaves emotions' root motivation unexplained—*why* one tends to experience negative emotions when one merely believes he was wronged. The obvious answer is that one also possesses a *desire* that the person

[8] *In Praise of Blame*, 94-95.
[9] Sher, *In Praise of Blame*, 79-80.
[10] Sher, *In Praise of Blame*, 78-84.

had not acted as he did.[11] This B/D pair then creates a disposition towards negative emotion that may commonly manifest in actual anger, etc.

Sher's view also makes sense of why the degree or severity of blame's negative emotions, if experienced, tends to roughly correspond to the level of the wrong; D's magnitude is based on the magnitude of the wrong as understood by B.

2. *Hostile Behavior*. Blame often results in hostile behavior—anything from "writing someone out of our will, sending someone a poison pen letter, sending someone poison, [or] urinating in someone else's flower bed."[12] Perhaps more typically, hostile behavior can take the form of yelling, severing ties, or storming out of the house. Like negative emotion, hostile behavior is a common result of blame but is not an essential component. We can blame without expressing ourselves in hostile ways. Similarly, hostile behavior is explained by Sher's pair, B providing the trigger and D the motivation for these expressions.

3. *Reproach*. Verbal or written reproach is a more specific class of blame's common manifestations. In either form, reproach has an essentially communicative dimension; it aims to be heard and understood by its target. When I reproach you for acting poorly or demonstrating a vicious character, I necessarily reproach you *for* something—the deed or trait. When I reproach you *for* something, I highlight the content of my belief that you acted wrongly or have a bad character, which again aligns with Sher's B/D pair.

---

[11] Wallace's Reactive Account, which is a further revision of a classical Strawsonian theory of blame (not that Strawson himself ever used the word), indicates an essential connection between a belief that someone has done wrong and the typical negative emotional responses: "on the reactive account, holding people responsible involves a susceptibility to a range of reactive emotions, so that to blame a person is to be subject to one of these reactive emotions, *because* of what the person has done" (1994, 75, emphasis added). There is clearly a connection between the belief of wrongdoing and the resulting emotions, but Sher's essential desire explains the motivational connection between the belief and the emotion—*why* we are disposed to feel resentful when wronged.

[12] Sher, *In Praise of Blame*, 94.

Here too we can observe how Sher's account stands above its rivals. A utilitarian account understands blame as simply reproach: "an expression of our disapproval."[13] Because being reproached is unpleasant, blame for the utilitarian is essentially "punishment light."[14] However, the utilitarian account fails to understand *why* reproach and directed hostile behavior are often effective as punishment light: because the reproach is presumably an expression of the speaker's genuine blaming attitude. Reproach only works on the assumption that there is a genuine desire behind it.

Likewise, blame cannot solely be a set of beliefs.[15] For one, the overwhelmingly common experience of blame involves *something* more than mere belief. Sher delivers a plausible answer: blame also involves a desire that provides blame with its punching power. Additionally, blame as mere belief(s) cannot explain blame's close association with anger and other negative emotions—why should I become *upset* by the mere belief that "someone's bad action has stained his soul or reduced his moral balance?"[16] Sher's account makes perfect sense of this connection; I often experience negative emotions because I have a desire that you not have done that wrong act.

4. *Apology*. Apology often results from self-blame or the recognition that one acted wrongly via other's blame. Like reproach, an apology is always offered *for* some specific wrong or trait; it makes little sense to apologize or blame (or forgive) in general.[17] This aligns well with Sher's B. When an apology is offered and accepted, both the wrongdoer and the offended party share a common belief that what occurred was wrong.

[13] Sher, *In Praise of Blame*, 74.
[14] Sher, *In Praise of Blame*, 73.
[15] Sher, *In Praise of Blame*, 75-76.
[16] Sher, *In Praise of Blame*, 78.
[17] Russell, "The who, the what, and the how of forgiveness," 3.

Blame Desiderata—Provided by Wallace (1994)

1. *More than Mere Evaluation*. While blame must involve evaluation, nearly all philosophers (and I dare say all normal folk) agree that blame must involve something more than mere belief. The entire point of blame-theorizing is to create a "useful distinction between *blaming* and *judging* blameworthy."[18] Sher follows suit by adding D to B, thus securing this desideratum.[19]

2. *Cannot Rule out Incompatibilism*. A theory of blame should not logically preclude incompatibilism; that would rig the debate so as simply define incompatibilism as wrong. Any theoretically satisfying account should be compatible with any possible position regarding moral responsibility.[20] Sher's B/D pair does not include any bloated metaphysical baggage on this front, and thus secures this second desideratum.[21]

3. *Backward-Looking*. Blame must be backward-looking (or at least not forward-looking).[22] It not only seems unfair to blame someone for what they might do, it seems rationally incoherent.[23] Our blame responds to violations done in the past or a vice demonstrated currently. This is why we do not blame people in general; rather, we blame people *for* certain wrongs. Sher's B and D both concern past or present violations and thus secures this desideratum.[24]

---

[18] Coates & Tognazzini, *Blame*, 9 (emphasis added).
[19] See Sher, *In Praise of Blame*, 75-78, 113.
[20] Wallace, *Responsibility and the Moral Sentiments*, 60.
[21] In fact, Sher's theory is so theoretically slim that several philosophers have argued it is underweight as far as blame theories go; see Pereboom (2013), Smith (2013), McGeer (2013), Franklin (2013). I will return to this idea in my attempt to constructively revise Sher's account.
[22] Wallace 67, 75.
[23] Even if we recall moments of particularly irrational rumination, wherein we seem to blame someone for something they *haven't* done but *could*, it seems we are still, in a rather insane way, blaming them for a wrong that will have *happened*.
[24] See Sher, *In Praise of Blame*, 101-103.

4. *Private Blame*. It should be possible to privately blame.[25] Accounts that define blame as simply expressions of reproach miss that we can blame others without them ever knowing. In fact, most blame might remain private. Sher's account, since it consists of a B/D pair, both of which are private mental states, clearly secures this desideratum.[26]

5. *Closely Linked to Public-Facing Expressions*. A satisfying theory should explain the tight connection between experiences of blame and the common expression of such attitudes.[27] While some accounts define blame in expressive terms, as we previously saw, this precludes the possibility of private blame. Sher's account makes perfect sense of this desideratum; B provides blame's content while D motivates its expression.[28]

Sher's account clearly secures our four blame explanandums and five desiderata. So, what *is* my problem? Why not just accept Sher's account? While I (1) affirm the importance of each blame explanandum and desiderata, (2) defend Sher's account as securing them all, and (3) find his account superior to other prominent theories, I believe Sher's account falls short because both lists are *incomplete*. In the next section, I will discuss two forgotten blame explanandums and one blame desideratum that must be included in our list of requirements. I will support these additions with both a theoretical defense and descriptions of common experience. The question will be whether Sher's account of blame can make sense of each new blame explanandum or desideratum.

Section 2: Excuses, Exemptions, and the Objective Attitude

[25] Wallace, *Responsibility and the Moral Sentiments*, 66-67, 74-75.
[26] See Sher, *In Praise of Blame*, 74, 113.
[27] Wallace, *Responsibility and the Moral Sentiments*, 67, 74-75.
[28] See Sher, *In Praise of Blame*, 105-106, 108-111, 113.

Excuses and exemptions constitute our additional blame explanandums. While neither are instances of blame themselves, like reproach, hostile behavior, and apology, they are common experiences directly connected to blame. While both have been widely discussed in the context of apt blame*worthiness*, they have not received such attention as a feature of a theory of blame itself.

Explanandum 5: Excuses

Excuses cancel blame. When an excusing plea is accepted by the blamer, he comes to believe the act itself was not one that warrants blame—that a condition obtains cancelling the blameworthiness of the offender.[29] When an excuse properly functions, it works to convince the blamer that the *act or trait* did or does not merit blame; a standard *internal* to the criteria of blameworthiness is not believed to be met.[30] P.F. Strawson provided many common excusing pleas: "'He didn't mean to', 'He hadn't realized', 'He didn't know'…'He couldn't help it'…'He was pushed', 'He had to do it', 'It was the only way', 'They left him no alternative', etc."[31] The central thought behind excuses is that one should not *blame* X because X had a valid excuse.

Some readers may have noticed the rather tortured emphasis on the blamer himself merely *believing* or *accepting* an excuse—I have emphasized this to focus on *excusing cases*, not *excusing conditions*. I am discussing what an excuse is and how it operates, not whether any given excuse is a normatively appropriate one. Thus, if an excusing plea is accepted and cancels blame, then an excuse has descriptively occurred, regardless of its normative appropriateness.

Excuses should clearly be added to our list of blame explanandums because they so clearly involve a cancellation of *blame*. Whatever blame is, excuses work to cancel it. I likely offer and

[29] Watson, "Responsibility and the Limits of Evil," 123.
[30] Strawson (2020) says the same regarding the criteria of the reactive attitudes, which is taken up at length in Wallace 1994, Ch3.
[31] Strawson, "Freedom and Resentment," 113-114.

receive several excuses a day; anything from a simple "*oops*"—indicating that an action was accidental—to "*Sorry…I've been under horrible stress from my tenure deadlines*."

Can Sher's B/D pair make sense of excuses? Yes. In fact, Sher explicitly discusses how his theory best explains excuses: "The obvious explanation here is that the belief upon whose truth [blame depends] is rendered false by each of the standard excusing cases."[32] This is exactly right; an excuse operates by changing the belief that someone performed a wrong action while the desire that the person not to have acted so may remain. But D without its supporting B is merely a preference, not blame; I would prefer for you not to have stepped on my hand even though I do not blame you for accidentally doing so. Excuses, then, are not an issue for Sher's account.

Explanandum 6: Exemptions

Exemptions also cancel blame, but they do so in a very different manner. While excuses cancel blame by rendering an *act* unworthy, exemptions cancel blame by rendering the *person* unworthy; the person himself is not seen as an apt target of blame due to some perceived feature about *him*. Exemptions cancel blame (or preclude its arising) because the other person is not seen as an apt candidate for blame—he is exempted from the moral community of people we stand in blaming relationships with.[33] Again, I will take pains to highlight this purely descriptive account of exemptions; an exemption occurs whenever it is the case that a blamer rescinds blame (or fails to ever blame) because he has shifted his view of the other and removed him from the moral community of those whom he is prepared to blame.

[32] Sher, *In Praise of Blame*, 118.
[33] Watson, "Responsibility and the Limits of Evil," 123, 130.

There may be two broad types of exemptions; one might exempt another from blame for *present-tense pragmatic* reasons, or one might exempt another because one comes to believe that *historical* information about his psychological composition renders him an inapt target of blame. In both cases, the person himself is exempted. In the former, he is exempted for the would-be blamer's practical concerns; in the latter because of some perceived historical fact(s) about him.

I will call the first type *pragmatic exemptions*—performed without respect to historical facts about the target's underlying psychological composition. I suspect these exemptions occur more often than we realize; we might exempt another simply because he is *unavailable* for dialogue—because he inhabits a disparate moral community. It seems we can (and do) exempt others for a wide array of practical reasons; we might foreswear blame in our role as professor when a sophomoric student defends an ethically dubious take on sexual consent, as a therapist hearing a patient's innermost demons, or to simply survive a Thanksgiving dinner with one's politically outspoken uncle. Below are (I hope) common examples of pragmatic exemptions.

> **1.** *Moving Forward*. Gerald and Lucy are college sweethearts. While Gerald is studying abroad, Lucy is out at a bar with some friends. After a couple drinks—just enough to be 'buzzed'—Lucy ends up sleeping with one of Gerald's classmates. In the morning, she instantly regrets her decision, realizing that she was foolishly motivated by loneliness. When Gerald returns, she tells him this, and understandably, Gerald is angry and aggrieved, and deeply blames Lucy for her infidelity. As they work through their issues, though, Gerald comes to relinquish his blame towards Lucy and commits to not blaming her for this transgression in the future. He wipes the slate clean,[34] promising to move forward and

[34] See Allais (2008). Wiping the slate clean for Allais is essential to forgiveness; it involves not holding an act against someone moving forward.

give the apologetic Lucy another chance. While Gerald has let go of his blame towards Lucy, he still thinks that what she did was wrong and wishes she hadn't done it.

2. *Left Cold*. Michael has been a judge in the Texas State Court system for 28 years. He has overseen every type of case imaginable. Earlier today, he sentenced a man to serve 10 years in prison for an armed robbery. While Michael, as an upstanding moral man, desires that the man had not ruined his own life and the lives of others, and clearly believes the armed robbery was wrong, he doesn't find himself *blaming* the criminal. In his 29th year, Michael simply doesn't muster up full-blown *blame* for each and every criminal he sentences.

To recap: *Moving Forward* cases tend to occur in the context of interpersonal relationships; one might move forward even though the other party demonstrates or demonstrated a vice. *Left Cold* cases tend to occur when one fails to blame another because he is left cold—he just can't muster or animate what he (and I) would call a full-blown blaming response. These cases might occur when one has been desensitized to wrongdoing, like Michael the judge has been; perhaps lawyers, therapists, schoolteachers, and historians often fit this bill.

Next, I will examine *objectifying exemptions*: cases where historical information about the other's psychological composition causes the blamer to view him as a merely unfortunate object, rendering him an inapt target of blame. In these cases, the exempted party not blamed because of some perceived psychological defect: he is deranged, warped, twisted; a product of bad upbringing, genetics, or disfigurement.[35] These exemptions involve viewing the other as an unfortunate outcome of prior causation, inviting an adoption of Strawson's *objective attitude*,

[35] Strawson, "Freedom and Resentment," 115.

which will constitute our sixth blame desideratum. Since the objective attitude and these historical debunking exemptions are so intimately related, I will discuss them both simultaneously.[36]

Desideratum 6: Incompatibility with the Objective Attitude

Humans tend to inhabit a natural participant stance, under which one is naturally reactive towards another and is prone to blame him should he commit wrong.[37] The objective attitude involves a suspension of this natural reactivity. It entails a change in view such that the would-be blamer comes to see the other as unworthy of "getting what he deserves" *because* he is seen as an unlucky byproduct of his environment, genetics, experiences, etc.[38] When one fully adopts an objective attitude towards another, that person is reduced to an *object* that can be reductively understood, perhaps "managed or handled or cured or trained."[39] He is seen as nothing more than a mechanistic object, like an automobile or plane. And while he may still be the most complicated, wonderful thing in the world, he is a *thing* nonetheless. Strawson and Watson highlighted that viewing another in these terms naturally invites an exclusion from one's moral community of those whom he stands in blaming relations towards. Coming to see someone with an understanding of how he came to be the twisted *thing* he is today eradicates one's impulse to blame him.[40]

Adopting the objective attitude drives our second subset of exemptions; as one "zooms out" and learns more about the history of another's current composition, one comes to see him in

[36] Our new explanandum of exemptions and new desideratum of blame's incompatibility with the objective attitude are directly related to each other—this mirrors the interconnected nature of the possibility of private blame, reproach, and hostile behavior, or of blame's backward-looking nature and the common emotional responses to wrongdoing.
[37] Strawson, "Freedom and Resentment," 116-118.
[38] Strawson, "Freedom and Resentment," 116. Sommers (2007) critically fails to mention that a reductive objection of the other is what drives one's belief that he is not worthy of blame.
[39] Strawson, "Freedom and Resentment,"116.
[40] See Hirji's discussion of this feature of the objective attitude in her defense of outrage anger (2022, 11-12).

a reductive light and naturally relinquishes blame.[41] Adopting the objective attitude can explain why one exempts another without having overtly pragmatic reasons for doing so.

Like excuses and exemptions, I believe adoptions of the objective attitude (and objective-attitude-driven exemptions) are robust and widespread phenomena; reading Strawson's description for the first time is an experience of more fully understanding what one has already done for years. And it should ring true that, whatever blame is, it is incompatible with fully adopting this stance.[42] Thus, our sixth desideratum: blame must be incompatible with the objective attitude. Below are two common examples of the objective attitude cancelling blame:

> **3.** *Non-Overlapping Moral Communities*: My hypothetical grandfather, Eugene, often expresses morally outdated views.[43] At the slightest provocation, he will spout off: homosexuality is a sin, a women's role is to live in servitude to her husband, etc. When I was a younger man, I would often intensely blame my grandfather when he spewed out such nonsense. But, as the years went by, I came to see him with something close to a completely objective attitude; I considered what his own childhood consisted of, of the backwards teachings instilled on his mother's knee, how he has been calcified by fear of the unfamiliar, and how his rural, cloistered life has hardly presented him with opportunities for reform. I realized we inhabited non-overlapping moral communities and came to fully understand how he ended up as he is—as an unfortunate byproduct of his own upbringing, environment, genetics, etc. For many years now, I have fully transcended

[41] Rosen, "Culpability and Ignorance," 73.
[42] While I cannot be sure, I would suspect that if pre- or non-philosophical folk were told about the objective attitude and asked to describe instances of their own adoption of the view, they would likely not describe cases wherein they saw themselves as "blaming" the target of such an attitude. Perhaps more empirical work awaits us on this front.
[43] Coincidentally, Rosen (2003, 66-69) uses an elder sexist named Smith as an example of someone who may be non-culpably ignorant for his unenlightened beliefs. My grandfather's name happens to be Eugene Smith.

blaming my grandfather. However, even from the objective stance, I retain the belief that his words and actions are morally wrong and desire that he possess a different character.

4. *Warped and Deranged*. Robert Harris committed unimaginably horrible crimes.[44] When first reading of his misdeeds, it seems obvious that Harris is both worthy of blame and is indeed *blamed* by first-time readers. However, his story is not so simple. After he was born nearly 3 months prematurely due to a kicking of his mother's stomach, Harris' childhood proceeded to meet or exceed the cruelty of his later actions. He was described by siblings as a gentle and kind soul who was beaten, raped, and emotionally abused by his parents and cellmates. He developed cognitive disabilities that were left untreated and tried to take his own life more than once before going on to kill and abuse others. The experience of reading Harris' story is emotionally challenging; I found myself first intensely blaming Harris for his heinous deeds, then coming to see him with a pity-toned objective attitude,[45] no longer blaming him for what he appeared to be warped and pushed towards doing. However, while I distinctly no longer *blamed* Harris, I retained the belief that his actions were morally horrendous and, empathizing with his victims, desired that he had not done his wicked deeds.

Section 3: Problems for the B/D Pair

[44] Harris' case is discussed at length in Watson (1987, 131-138).
[45] Strawson, "Freedom and Resentment," 116.

I believe Sher's account of blame calls for revision because it (1) cannot make sense of exemptions and (2) proposes an account of blame that is perfectly compatible with the objective attitude. Our new explanandum and desideratum raise problems for Sher's theory of blame.

The two-sentence version of my critique is as follows: Exemptions and adoptions of the objective attitude cancel or preclude blame while Sher's B/D pair can congruently remain present. Therefore, it appears as though blame must involve more than belief and desire.

Experiences of exempting another, like those presented in our four cases, should be highly relatable. I have often cancelled my initial blame, preemptively foresworn blame, or naturally failed to ever blame another. But in all four cases, while blame is rescinded, the belief that wrong was done remains. Harris' murders and Eugene's expressions of sexism were still wrong, as were the robber's theft and Lucy's infidelity. Additionally, Harris and Eugene clearly retain a vicious character.

Nor can exemptions be located in a change of desire. In all four cases, while blame is withdrawn, the desire that the malefactor had not committed the wrong remains. I still desired that Harris had not murdered, and that Eugene had a different character, because the world would be a better place if either were different. The judge retained his desire that the robber had not robbed, and Gerald clearly wished for Lucy never to have cheated. So, while their exemptions cancelled blame, that cancellation cannot be located in a change of belief or desire. In exempting cases, something other than B or D is lost when one goes from *blaming* to *not blaming*. This suggests that blame must have some additional element to fill this gap.

Perhaps Sher would reply that my cases are instances wherein these 'villains' are *not* exempted because they are all in fact still deeply blameworthy.[46] While this could well be true, recall that my argument does not concern whether any given person is blame*worthy*, but rather, only if one simply *does* exempt these people from blame. My argument is within the descriptive domain, not the normative. And allow me to confess that I do genuinely exempt both Eugene and Harris, and I have been in similar situations to both Gerald and Michael before. I suspect that readers can draw on their own experiences of exempting another while B and D remain. This suggests that there must be something else to blame.

Before we arrive at our solution, though, we should examine if Sher's account is sufficiently incompatible with the objective attitude. Recall that adopting the objective stance entails seeing its subject as not deserving what he was due *because* he is viewed as an unfortunate mechanistic object or outcome. Whatever blame is, it should be incompatible with this stance.

However, Sher's B/D pair is perfectly compatible. One can have B and D while also believing that their target does not deserve what is coming to him and while reductively seeing him as a mere outcome. Harris and Eugene are clear examples.[47] I came to believe that neither deserved what they were due,[48] and came to believe that because I learned or realized more about their history. These exemptions were driven by an adoption of the objective attitude which was itself driven by learning more about how either came to be molded into who they are. Yet I retained the belief that they acted wrongly and had vicious characters, and I desired both that they had acted differently then and were different now.

[46] Sher, personal communication.
[47] Michael & Gerald are exempted for *pragmatic* reasons outside the scope of objective-attitude-driven exemptions.
[48] My blame as well as any purely retributivist punishment.

Perhaps Sher would reply that his account of blame is not in fact compatible with the objective attitude because the belief that one acted wrongly requires viewing him as an agent, which is itself incompatible with the objective attitude. Thus, B would indeed change if the objective attitude was adopted, making Sher's theory properly incompatible.[49]

I see two issues with this reply. The first is that believing that someone did wrong does not appear to *require* viewing him as agential. As Derk Pereboom argues, even assuming that ought implies an agential 'can,' it would not follow that murder was not morally wrong. Even from an incompatibilist's objective stance, "one can replace occurrences of 'you ought to do x' with 'it would be right for you to do x,' or with 'it would be a good thing for you to do x.'"[50] And insofar as a wrong act is both clearly wrong and clearly attributable to a certain person, there is no conflict in holding a belief that one acted wrongly while inhabiting the objective stance.

The second issue follows a similar line. If B requires seeing the target of blame as acting wrongly, and 'acting wrongly' is spelled out in terms of being an agent who can freely act without prior 'objective' determination, this seems to violate our second blame desideratum: that blame should not definitionally rule out incompatibilism. If acting wrongly *qua agent* requires that he cannot be viewed in reductively objective terms (as an outcome of prior causation), then this appears to be defining incompatibilism out of the picture. To remedy this issue, B need not require seeing the wrongdoer as an agent—how I have conceptualized the belief. It is the belief that the person did wrong, even if they were warped or deformed or otherwise pushed into their villainous role, as was the case with Harris and Eugene.[51]

---

[49] Sher, personal communication.
[50] Pereboom, "Determinism al dente," 36.
[51] Sher might again reply that I've made the same mistake with exemptions—that I baked in conclusions about incompatibilism being right or wrong because I suggested that taking the objective stance sees someone as an outcome of causation, which cancels blame, which just means incompatibilism is right. But I have not done this,

But while we have seen that exemptions cannot change B, Sher's objection is clearly on to something; when we shift to an objective stance, we are clearly seeing the person in different terms than we naturally do. When we exempt someone from blame, something about how we relate to him changes, but I have argued that this change cannot be found in B or D.

So here lies my position: I cannot accept Sher's view because it cannot make sense of exemptions and is wholly compatible with the objective attitude. Yet, I cannot reject Sher's view because his B/D pair explains the original blame explanandums and desiderata (plus excuses) better than any other account of blame. The most satisfying move would be to revise Sher's theory so that it could better understand our common phenomena. However, proposing additional elements to Sher's account will bloat the theory and render it less theoretically satisfying. It would be wonderful if only one additional element could solve all our issues in a single stroke. Luckily, I believe there is such a solution.

Section 4: The *BAD* Theory of Blame

Sher's theory of blame can be successfully revised by changing just one word of his account: instead of a B/D pair that *disposes* one towards negative affect, blame should be understood as a B/D pair that *does* create negative affect. Thus, the additional element to blame is negative affect, creating the *BAD* theory of blame: belief, desire, affect. I will argue that these three elements are necessary and sufficient for understanding blame as a B/A/D trio.[52] Adding negative affect to Sher's more elegant account (1) solves the problems raised by exemptions, (2) makes

because, following Wolf (1980), we do not exempt people in objectifying cases because we come to believe determinism is true; we do so because of *how* we believe the person was determined—unfortunately or unluckily!

[52] I am aligning with many philosophers who see directed negative emotion as necessary for blame. See Bell (2013, 265-266), Wallace (1994), Strawson (2020), Franklin (2013), Rosen (2003, 66), Wolf (2011).

blame properly incompatible with the objective attitude, and (3) better aligns with a wide array of common experiences regarding blame.

Exemptions and Objective Attitude Incompatibility

The *BAD* theory of blame allows us to understand exemptions. Recall that exemptions changed something about the would-be blamer's response (from blame to not-blame), but that change could not be found in either B or D. The *BAD* view of blame explains that change by the disappearance of directed negative affect.

Recall our four cases; in each instance, blame was cancelled while B and D remained. I now suggest that what explains *blame's* disappearance is the loss of negative affect directed at the wrongdoer. In Gerald's case, by moving forward with Lucy, he let go of his rage, disappointment, and indignation. Michael failed to ever resent or despise the robber because he had become so accustomed to dealing with wrongdoers. While he still believed their acts were wrong and desired that they had not occurred (for the betterment of the world), he did not find himself *blaming* petty thieves who crossed his courtroom. When I relinquished blame against Eugene, I did not come to believe that his actions or character were justified, nor did I lose my desire for him to be a different person; rather, I relinquished my disgust, contempt, anger, and disdain. Finally, when learning of Harris' tortured upbringing which twisted him into the monster he became as an adult, my directed negative affect dissipated; I was no longer contemptuous and enraged while I certainly retained my B and D regarding his actions.[53]

This raises an important clarification. I stated that exemptions—in the last two cases driven by an adoption of the objective attitude—cancel blame by removing one's *directed* negative

[53] Mirroring Watson's own reaction (1987, 137-139).

emotions. *Directed* affect broadly refers to that subset of Strawson's reactive attitudes which are both negative and connected to blame: anger, resentment, indignation, contempt, etc. The directed nature of blame's negative affect is critical, since the objective attitude can itself be emotionally toned, but not in a *directed* manner; "it may include repulsion or fear, it may include pity…but it cannot include the range of reactive feelings and attitudes which belong to involvement or participation with others in inter-personal human relationships."[54] Watson raises this point nicely when suggesting that learning of Harris' background invites us to feel an 'objective' sense of fear, pity, and sorrow for his utter misfortune.[55] But these emotions are not directed *at* Harris—they are a way we feel *about* him.[56] This also explains why A stands for *affect* rather than *anger*. Anger alone might be too narrow, even if we include its close relatives. Sometimes blame involves emotions that might not be species of anger, such as disgust, contempt, disdain, etc.

Thus, the *BAD* theory makes blame properly incompatible with the objective attitude. As Strawson and Watson suggested, the objective attitude invites one to relinquish directed (reactive) negative affect. While it is possible to, in a fit of blind rage, become angry *at* your car for not starting, when you come to view it properly—as a collection of normally interworking parts that are not working for some reason now—it is impossible to remain angry *at* the car, while you may remain irritated or frustrated *with* the situation.[57] Likewise, when I adopt the objective attitude towards Eugene and Harris, and come to see them as *malfunctioning* or *deformed* rather than *villainous*, I am invited to relinquish my anger, disgust, contempt, and resentment *towards* them while I may well retain my fear *of* them, pity, and sorrow *that* they are so warped.

[54] Strawson, "Freedom and Resentment," 116.
[55] Watson, "Responsibility and the Limits of Evil," 137-139.
[56] 'Directed' does not mean that the emotion is expressed; it simply means that one is angry *at* or *towards* another.
[57] This distinction shows up in Hirji differences between reform anger and outrage anger.

Giving Bargle His Due

Bargle suggested that if a theory diverges too far from common experience, this should give us reason to reconsider its plausibility. I have spent much of this paper attempting to realign Sher's theory of blame from a perspective amendable to both Bargle and Argle, arguing that exempting others and adopting the objective attitude have deep connections with more theoretical values. But here it is it time to give Bargle his day; the *BAD* theory of blame not only remedies theoretic desiderata that seduces analytic philosophers; it also better aligns with common sense and common experience.

Recall Lucy's infidelity. When Gerald decides to move forward with their relationship and wipe the slate clean, he lets go of his *blame* towards Lucy. But because he may well retain his belief that what she did was wrong and his desire that she had not done it, Sher's theory states that he is continuing to blame her, even while moving forward with the relationship. On Sher's account, Gerald has merely given up his *anger*, while on my account, he has given up his *blame*. I will admit that this consequence of Sher's theory strikes me as implausible: Sher has to say that the couple moves forward with the relationship, rebuilding love and trust, all the while Gerald continues to actively blame Lucy. When he forgives her, builds trust in her, and falls back in love with her—while retaining his B & D (and why would he not retain them?)—he does so while altogether blaming Lucy. This is simply antithetical to my own experience.

Sher's account also makes blame far too prevalent; we are blaming others all the time if blame sufficiently consists of a B/D pair. I believe people do wrong all the time, and as a decently moral person who is committed to advancing moral aims, I have the desire that (at least) much of the world's wrongdoing be otherwise. So then, it seems as though whenever I have the additional belief that someone did wrong or demonstrates a bad character trait, I am 'full-on blaming' them.

Again, this strikes me as disconnected from how I actually go about blaming others. The *BAD* theory of blame solves this issue; while I may widely have Sher's B/D pair, only some subset of those instances creates directed negative affect. Thus, blame is still quite common, but less ubiquitous than Sher's account suggests.

The *BAD* theory of blame thus creates the following conceptual ramp-up:

1. *Cognitive Evaluations*: Merely believing X to be wrong.[58]

2. *Conative Evaluations*: Believing X to be wrong and desiring that X had not occurred (Sher's B/D pair).

3. *Blame*: Believing X to be wrong, desiring that X had not occurred, and experiencing negative blaming affect directed towards the wrongdoer.

These distinctions are more than theoretical fodder. The BAD theory of blame sees an important difference between all three states, while Sher's account lumps together both 2 and 3 under blame. On Sher's model, possessing a B/D pair while remaining completely affectless is just as much blaming as when someone's B/D pair leads them to be enraged and overwhelmed with contempt. Those two experiences seem quite different, though, and the *BAD* theory respects this distinction.

Finally, the *BAD* theory of blame explains why only some of Strawson's reactive attitudes fit within the purview of blame. Requiring blame to include *directed*, *negative*, *blame-associated* affect makes sense of cases wherein one has a B/D pair, and directed affect, but does not intuitively *blame*. Angela Smith speaks of a mother whose son was justly convicted of murder. The mother believes that he acted wrongly and desires that he had not committed the crime, but, seeing her

[58] This can include judging a state of affairs as desirable without desiring it. See Sher, *In Praise of Blame*, 103.

only son taken away in handcuffs, feels only remorse, despair, grief, and unconditional love.[59] The mother does not appear to be blaming her son—indeed, "some of these reactions…seem to be the opposite of blame."[60] While this perfectly aligns with the *BAD* theory of blame, Sher's account stated that blame sufficiently consisted of the B/D pair. Thus, one blames whenever this pair obtains. And while Sher's account can explain the typical connection between the B/D pair and typical blaming affect (anger, resentment, etc.), his account produces the unintuitive conclusion that this mother is indeed blaming her child when she experiences remorse, despair, grief, and unconditional love. Whenever the B/D pair obtains, no matter how un-blame-like the additional emotion, one is said to blame the other. Following Smith's lead, this conclusion is too askew with common experience to secure a satisfying account of blame. Luckily, the *BAD* theory of blame, taking its cues from Sher's original account, can solve this mismatch.

Section 5: Objections and Clarifications

Objection 1: "Why bother with D at all?—It's not obviously doing any work in the *BAD* account of blame. If you add A, D's motivational role becomes redundant, making you appear to just reembrace Wallace's Reactive Account of blame with all its problems."

Reply: Not so! The *BAD* theory does require D's presence because it explains why B could lead to A. Remember that the *BAD* account takes its cues from Sher; the belief that an act was wrong and the desire that it had not been so explain *why* one would become angry, resentful, etc. in response to a wrong. The desire provides motivational force that creates blaming affect as well

[59] Smith does not include unconditional love in her example.
[60] Smith, "Moral Blame and Moral Protest," 35. While Smith also raises this example against Sher's account of blame, I do not follow her line that Sher's "core desire-belief pair is neither necessary nor sufficient" (35). I believe the B/D pair is necessary but insufficient for blame.

as often-accompanying hostile behavior and reproach.[61] B gives blame its content, D provides the motivation, and A, for my account, provides the emotional reactivity. This is why my account explicitly rejects Wallace's Reactive Account or Bell's Hostile Attitude Account;[62] of course blame is more than belief, and I align with Wallace and Bell in seeing blame as requiring negative affect, but the connection between the two is left unexplained without D.[63]

Objection 2: "You've acted as though all the common experiences and intuitions fall on your side of the argument. But there are some on Sher's side too. Don't you blame Mussolini for what he did, even though you don't currently have any negative affect directed towards him? I sense that you'd have a hard time saying 'no'—isn't that a point in Sher's favor? And if there's intuitions on both sides, why should the *BAD* view of blame win out?"

Reply: This is a powerful objection, since much of my argument has relied on the *BAD* account's ability to incorporate more of our common intuitions. And the objector is right—even though I don't have any affect towards him, I do find myself having a hard time saying 'I don't blame Mussolini.' However, I think this difficulty can be explained away. In fact, I was asked this very question when discussing this paper, and I did feel comfortable stating that I didn't blame Mussolini, while immediately adding that I had a certain belief that what he did was wrong and strongly desired that he had not done what he did. I think that 'Mussolini cases' only appear to support Sher's account because we understandably have a hard time admitting that we do not blame Mussolini—we are, I suggest, inchoately concerned that 'I don't blame him' could indicate that we do not possess B or D, while in reality, we merely lack A. In general, when one fails to blame another, it could be because one fails to obtain the B, D, or A. Failing to believe that what

[61] Sher, *In Praise of Blame*, 98-103.
[62] Bell, "The Standing to Blame: A Critique," 265. & Wallace, *Responsibility and the Moral Sentiments*, 75.
[63] Likewise, D and A without B is merely preference (non-normative).

Mussolini did was wrong or failing to desire that he had not done what he did are two possible explanations for the statement that one does not blame him. But apprehension and concern for being clear on *which* aspect of blame one lacks does not mean that one does indeed blame Mussolini. So, I believe I can deflate the seemingly strong impulse supporting Sher's view; when I have trouble saying that I do not blame Mussolini, it is not because I do in fact blame him, but rather, because I am afraid that my audience will mistakenly believe that I therefore lack B or D.

Conclusions

While this investigation has primarily concerned the nature of blame, I want to briefly remark on some broader normative implications that follow from adopting the *BAD* theory of blame. First, adding directed negative affect makes blame a somewhat weightier moral project. When one becomes angry, resentful, etc. as a feature of their blame, this likely creates more 'leakage' into the outside world. While *BAD* blame can still be private (unexpressed), adding negative emotion makes one's blame more likely to manifest in *some* way; perhaps it makes the blamer slower to aid the 'blamee,' less likely to treat him fairly in an unrelated context, or quicker to rub salt in his wounds later on. While a B/D pair could still create these manifestations, adding A seems to make them much more likely. Thus, the *BAD* theory of blame seems to more directly call for normative concern regarding only *blaming* when the target is *blameworthy*—because *BAD* blame is more likely to manifest in external outcomes. Relatedly, adding affect explains why blame so often overshoots its mark. When we are angry, we tend to both overinflate the significance of actual wrongs and detect wrongs that are only imagined.[64] Finally, I believe the *BAD* theory can

[64] Pettigrove, "Meekness and Moral Anger," 361-365.

retain blame's essential connection to objective normativity.[65] Requiring a well-grounded belief that one's target has indeed done wrong explains why it would be wrong to become angry or resentful towards him if he was not blame*worthy*. As with Sher's original account, if B remains directly linked to D (and now A); anger is inappropriate if formed on an unfounded belief.

This concludes my offer to Sher's account. I propose a small theoretical addition—adding directed negative affect as a necessary condition of blame—in exchange for both theoretical payoff—the ability to understand exemptions and render blame incompatible with the objective attitude—and better fit with common experience. I believe this trade is a fair one, retaining both Argle and Bargle's beloved values. The *BAD* theory of blame can be accepted by philosophers and common folk alike.

[65] Sher specifies morality, but I take a more pluralistic view of what normative domains might aptly animate blame.

Works Cited

Allais, Lucy (2008). Wiping the Slate clean: The heart of forgiveness. *Philosophy and Public Affairs* 36 (1):33–68.

Bell, Macalester. (2013). The Standing to Blame: A Critique. In D. Justin Coates & Neal A. Tognazzini (eds.), *Blame: Its Nature and Norms*. Oxford University Press. pp. 263-281.

Coates, D. Justin & Tognazzini, Neal A. (2013). The Contours of Blame. In D. Justin Coates & Neal A. Tognazzini (eds.), *Blame: Its Nature and Norms*. Oxford University Press. pp. 3-26.

Franklin, Christopher Evan (2013). Valuing blame. In D. Justin Coates & Neal A. Tognazzini (eds.), *Blame: Its Nature and Norms*. Oxford University Press.

Lewis, David K. & Lewis, Stephanie (1970). Holes. *Australasian Journal of Philosophy* 48 (2):206 – 212.

Hirji, Sukaina (2022). Outrage and the Bounds of Empathy. *Philosophers' Imprint* 22 (16).

McGeer, Victoria. (2013). Civilizing Blame. In D. Justin Coates & Neal A. Tognazzini (eds.), *Blame: Its Nature and Norms*. Oxford University Press. pp. 162-188.

Pereboom, Derk (1995). Determinism al dente. Noûs 29 (1):21-45.

Pereboom, Derk. (2013). Free Will Skepticism, Blame, and Obligation. In D. Justin Coates & Neal A. Tognazzini (eds.), *Blame: Its Nature and Norms*. Oxford University Press. pp. 189-206.

Pettigrove, Glen (2012). Meekness and 'Moral' Anger. *Ethics* 122 (2):341-370.

Rosen, Gideon (2003). Culpability and Ignorance. *Proceedings of the Aristotelian Society* 103 (1):61-84.

Russell, Luke (2020). The who, the what, and the how of forgiveness. *Philosophy Compass* 15 (3).

Sher, George (2006). *In Praise of Blame*. New York, US: OUP USA.

Smith, Angela. (2013) Moral Blame and Moral Protest. In D. Justin Coates & Neal A. Tognazzini (eds.), *Blame: Its Nature and Norms*. Oxford University Press. pp. 27-48.

Sommers, Tamler (2007). The objective attitude. *Philosophical Quarterly* 57 (228):321–341.

Strawson, P.F. (1962). "Freedom and Resentment. In P. Hieronymi. *Freedom, Resentment, and the Metaphysics of Morals*, Princeton: Princeton University Press, 2020, pp. 107-134.

Wallace, R. Jay (1994). *Responsibility and the Moral Sentiments*. Harvard University Press.

Watson, Gary (1987). 4. Responsibility and the Limits of Evil: Variations on a Strawsonian Theme. In John Martin Fischer & Mark Ravizza (eds.), *Perspectives on Moral Responsibility.* Cornell University Press. pp. 119-148.

Wolf, Susan. (1980). Asymmetrical Freedom. *The Journal of Philosophy*, *77*(3), 151–166.

Wolf, Susan, (2011). "Blame, Italian Style", in Wallace, Kumar, and Freeman (eds.), *Reasons and Recognition: Essay on the Philosophy of T. M. Scanlon*, New York: Oxford University Press, pp. 332–347.