LOVING YOUR ENEMY

Austen McDougal (austenm@princeton.edu)
Princeton University

Forthcoming in Oxford Studies in Philosophy of Religion.
Penultimate draft. Please cite published version when available.

Abstract: This paper begins by bringing love and hate into tension via the ideal that you ought to love your enemy. The trouble with loving your enemy is that they may seem to merit hate instead, especially in cases of serious injustice. I develop this simple thought into a challenge for loving your enemy: that you cannot be required to do what makes no sense to you. This challenge is not adequately met by extant explanations for why you ought to love your enemy within the Christian tradition and its heirs, which tend to give reasons that are either insufficient or else instrumentalize love. The second half of the paper presents a solution with a very different shape to the challenge. I argue that love may still be fitting even when your enemy fails to merit love, notwithstanding the contemporary orthodoxy about fittingness. What makes it fitting to love your enemy depends on the fact (if it is one) that you yourself have received unmerited love. You may thereby have reason to exclude the issue of merit from consideration; what is more, you have reason to love out of gratitude for the unmerited love that you yourself received.

"You have heard that it was said, 'Love your neighbor and hate your enemy.' But I tell you, love your enemies and pray for those who persecute you, that you may be children of your Father in heaven. He causes his sun to rise on the evil and the good, and sends rain on the righteous and the unrighteous. If you love those who love you, what reward will you get? Are not even the tax collectors doing that? And if you greet only your own people, what are you doing more than others? Do not even pagans do that? Be perfect, therefore, as your heavenly Father is perfect." (Matthew 5:43-48)

1. Introduction

If there were a ranking of the most challenging ethical requirements, near the top would be

(1) LOVE YOUR ENEMY: you ought to love your enemy.

It should go without saying that love of enemy runs counter to natural inclinations. Indeed, that might explain why LOVE YOUR ENEMY, despite appearing repeatedly throughout the Christian

New Testament, tends to be conveniently ignored by philosophers and non-philosophers alike (including in cultures supposedly saturated with Christian thought).[1] By contrast, the norm that you ought to love your neighbor—someone who may be a stranger but is not necessarily an enemy—is more comfortable to think about.

The difficulty with loving your enemy in particular is that they can strike you as, in a word, hateful: that is, they seem to you to merit hate.[2] Presumably, that is why the *endoxa* (rejected by Christ in the epigraph) tells you to hate your enemy. Hate and love, as I will explain in much more detail below, are opposing orientations toward a person, where reasons for one attitude count against the other. Insofar as your enemy seems hateful to you, then, you appear to lack sufficient reason on the basis of which you might love them. Love would make no sense—and you cannot be required to do what makes no sense to you.

Although the ideas underlying this challenge are nothing new, I aim to articulate a sharper version of it, drawing on contemporary work in analytic ethics and moral psychology. Later, I double back to show how loving your enemy can be reasonable, after all, by rendering precise a natural thought that has gone missing from today's analytic scene. The importance of these issues is not limited to intramural debates for religious ethics. Religious or not, love of enemies is no small matter, especially given the much-decried polarization of our times. And—if that is not enough—the nature of love, the possible rational bases for attitudes more generally, and the metaethics of fittingness are all at stake.

[1] I, like many of my interlocutors, am heavily influenced by the Christian tradition in this paper. That should not be taken to imply that Christianity has an exclusive interest in these issues.

[2] "Hateful," according to the OED (https://www.oed.com/dictionary/hateful_adj?tab=meaning_and_use), can refer either to the property of meriting hate or to the vicious disposition of being hate-filled toward others. I use it exclusively in the former sense, which parallels my use of "lovable" in the sense of meriting love. Of course, there is overlap in extension between the two senses of "hateful" since the clearest examples of people who seem to merit hate are often those who are vicious and hate-filled.

2. Sharpening the Challenge

The first half of the paper will develop the challenge previewed above via the following set of plausible claims that jointly rule out LOVE YOUR ENEMY:

(2) YOUR ENEMY IS HATEFUL: you take your enemy to merit hate.

(3) THE HATEFUL ARE UNLOVABLE: if you take some person *E* to merit hate, then you take *E* not to merit love.

(4) UNLOVABILITY IS DECISIVE: if you take *E* not to merit love, then you take there to be decisive reason against loving *E*.

(5) DECISIVENESS PRECLUDES REFLECTIVE LOVE: if you take there to be decisive reason against loving *E*, then you cannot reflectively love *E*.

(6) OUGHT IMPLIES REFLECTIVELY CAN: you ought to φ only if you can reflectively φ.

(C)Therefore, it is false that you ought to love your enemy.

We might say that the challenge for LOVE YOUR ENEMY, beyond being difficult, is that it would be positively irrational. For (as I will explain), taking a potential response to be merited or unmerited involves taking there to be normative reasons that serve as your basis for or against the response. Of course, you may sometimes judge others too harshly; it would be better to view things more judiciously. In any case, whenever you see someone as an enemy who merits hate (accurately or not), the challenge is thereby present. This is why (2)-(5) are stated in subjective terms of "taking." But also, some enemies really are as outrageous as they seem. Let us try to keep those cases in mind in order to take the challenge most seriously.

Philosophers ranging from Aquinas to Kierkegaard to Martin Luther King Jr. have sought to defend LOVE YOUR ENEMY. Yet even readers who find the ideal unattractive should note that there remains a philosophically important task: to understand whether and how it might be

intelligible for some people to succeed in loving their enemy, as indeed seems to be the case. The first half of the paper will now explicate (2)-(6) in order to show that there is a strong presumptive challenge against the possibility of loving your enemy as an ethical ideal. This challenge that I develop is not adequately met by extant explanations for why you ought to love. It is instructive to start with those familiar approaches, which are misguided in my view.

In the second half of the paper, I will offer a more promising solution to the challenge. Ultimately, I do reject UNLOVABILITY IS DECISIVE. But—to motivate the challenge at the outset—it is worth observing that UNLOVABILITY IS DECISIVE is at least *defeasibly* true: if you take it that someone fails to merit love, then you take there to be very strong reason against loving them. Strong reason is not necessarily decisive, of course, but the challenge is to show why not. Moreover, unlovability must fail to be decisive in a systematic way if LOVE YOUR ENEMY is to survive as something like a universal norm, even in the face of those who appear profoundly evil. Such enemies can fail to merit love, *pace* those who have deemed human nature a sufficient ground for love. Neither do prudential nor moralistic reasons for loving your enemy, however powerful, resolve the challenge. Rather, the richest source of reasons for loving your enemy lies in your background experiences with other relationships.

3. Your Hateful Enemy

There is little love lost on hate in philosophy today. In contrast to the discipline's eagerness to talk about love, the suggestion that a person could merit hate will sound especially jarring to our "cool, contemporary style" (Strawson 1962: 192). It is dubious simply to deny that an enemy might merit hate, however; it strains belief to think that you have never felt this way in your own life:

(2) YOUR ENEMY IS HATEFUL: you take your enemy to merit hate.

For those who want to resist, let us consider the challenge of loving your enemy from the position

of Christ's immediate audience. For those ancient Jews, the notion of an enemy would have quickly summoned to mind their Roman occupiers. In addition to suffering under Roman imperialism in general, many Jews would likely have had encounters with particular Roman soldiers, say, who exhibited great cruelty.[3]

An individual like Ben-Hur of the eponymous 1959 film offers a useful perspective for us to inhabit. In the story, Ben-Hur is betrayed by his erstwhile friend Messala, who has become a Roman general. Despite Messala knowing that Ben-Hur and his family are not guilty of a freak injury that befalls the Roman governor outside their home, Messala unjustly condemns Ben-Hur to the galleys, imprisons Ben-Hur's mother and sister, and appears to relish the fact that they are living out their days with leprosy as a result. Messala has no excuse for what he does. He intentionally deploys his privilege to harm Ben-Hur and his loved ones at every turn. Not only would it be apt to describe Messala as Ben-Hur's enemy, I claim that Ben-Hur would not be mistaken to see Messala as meriting hate.[4]

Granted, hate is not a nice attitude. It involves setting yourself against another in such a way that, *inter alia*, you desire and would be pleased by what is bad for them. Notice, though, how a similar thought underlies the natural retributive idea that someone who behaves in an egregiously bad way deserves to suffer for it. Hate is the more personal, affect-laden analogue, which can be merited especially when the offender behaves badly toward *you* and *yours*. Setting aside the question of whether it is desirable to be in such a state, hate seems reasonable—even to us at a distance—as a response to personal injustice. From Ben-Hur's own perspective, hate would surely

[3] In contrast to the notion of enmity that tends to appear in the context of just war theory, LOVE YOUR ENEMY is more personally involved. Both Greek and Latin offer a distinction between *polemios* / *hostis* (political foe) and *ekhthros* / *inimicus* (personal enemy). The sense appearing in Matthew and related passages is the latter, personal sense (Schwab 1987). If Christ's words brought the Romans to mind for his original audience, that is not ultimately because they were political enemies, I would suggest, but rather reveals that the political was personal for them, as it often is.

[4] Messala himself seems to recognize this in his own twisted way. Dying, he tells Ben-Hur, "There's enough of a man still left here for you to hate."

be positively gripping.

Not every case of enmity is like this, of course. Natural language allows a variety of conflicts to be described as "enmity," ranging from the extreme, like Ben-Hur's, to the relatively petty. Rivals in professional and social settings are often characterized as bitter enemies: e.g., romantic adversaries vying for the same partner, or sports competitors and their fans. There can even be enmity without hatefulness in grave matters of life and death. You could naturally describe a soldier on the opposing side as your enemy, for example, without any implication that you take them to merit hate. It need not be personal.

That said, any situation that pits your interests against someone else's risks you slipping into the thought that they merit hate. However unreasonable that may in fact be, it too suffices for the challenge to arise. I focus on cases in which your enemy really does seem to merit hate, even to us as impartial observers, in order to allow the challenge its natural force. Strictly speaking, though, the challenge is essentially psychological and therefore first-personal, as I explain further below: how can there be sufficient reason *for you, on the basis of which* to love your enemy? In order to emphasize this point, I frame the issue in terms of loving *your* enemy.

We might more readily agree from a distance that someone merits hate while continuing to try not to think in these terms about our contemporaries. Yet I suspect that many (if not most) of us continue to privately experience certain people as hateful. When I speak of your enemy, I invite you to substitute such people: the bully who enjoyed humiliating you; the scumbag who assaulted your friend; the demagogue who opposes your human rights; the extremist bent on destroying your way of life. In these cases, your enemy gives you powerful reasons to hate them.[5]

[5] I do not want to minimize the importance of epistemic humility. Some disputes are messy, and we would do well to listen to those urging a different side to the story. Still, some cases will strike us as clear-cut—and some indeed are.

4. Hate Rationally Contrary to Love

No attitude is an island, and hate shares an especially tight connection with love:

(3) THE HATEFUL ARE UNLOVABLE: if you take some person *E* to merit hate, then you take *E* not to merit love.

This connection holds in virtue of the psychological profiles of love and hate.[6] Short of a complete analysis, I make the following minimal assumptions. Love and hate are contrary orientations toward a particular person, constituted at least in part by syndromes of conative and affective attitudes such as desires and emotions regarding that person.[7] It is natural to understand hate and love as opposites in this sense: if hate manifests in a valenced response toward some outcome, love would manifest in the oppositely valenced response.[8] For instance, love is positively oriented toward the wellbeing of its object whereas hate is negatively oriented. Among other things, you would be motivated in favor of and pleased by what is good for your beloved, but motivated against and disappointed by what is good for someone you hate.[9]

It must be emphasized that these attitudes are not instrumental. Love held merely as a means to some further end is no love at all. Merely doing someone a good turn, where you can do so for a variety of reasons, is insufficient for the purpose of complying with LOVE YOUR ENEMY. As a colorful illustration, consider a different biblical injunction to act so as to help your enemy,

[6] It is plausible to think that there are various species of love, such as the four kinds described by Lewis (1960). However, I will assume that there is a common core to all such loves that we can bring under the heading of the sort of orientation described here. The challenge for loving your enemy is meant in this relatively thin, generic sense (although it may turn out that the solution bears on which kind of love is applicable toward enemies—see note 19).

[7] Although it may be controversial which of these are essential to love, what I say in this paragraph is meant to be uncontentious as a rough characterization. For some key views that I think would allow that love typically involves both the affective and the conative, see Kolodny (2003), Frankfurt (2004), and Ebels-Duggan (2008).

[8] People often like to say that "hate isn't the opposite of love, indifference is." Following Berker's (2024) helpful analysis of opposition, we can allow that indifference is the *privative* opposite and hate the *polar* opposite of love.

[9] Cf. Westlund (2009: 528): "As Butler would hasten to point out, one certainly cannot bear another goodwill if one goes beyond resentment to malice or hatred, which seem conceptually to involve wishing another ill, and to represent the other as meriting the opposite of care."

one that precedes Christ:

> If your enemy is hungry, give him food to eat;
> if he is thirsty, give him water to drink.
> In doing this, you will heap burning coals on his head,
> and the Lord will reward you. (Proverbs 25:21-22)

In this Old Testament proverb, you are told to care for your enemy, but not necessarily on the basis of caring *about* your enemy. In fact, feeding your enemy in order to heap burning coals on his head indicates that you still hate him. Similarly, Christ builds up to the injunction to love your enemy by starting with a weaker directive:

> You have heard that it was said, 'Eye for eye, and tooth for tooth.' But I tell you, do not resist an evil person. If anyone slaps you on the right cheek, turn to them the other cheek also. And if anyone wants to sue you and take your shirt, hand over your coat as well. If anyone forces you to go one mile, go with them two miles. Give to the one who asks you, and do not turn away from the one who wants to borrow from you. (Matthew 5:38-42)

All that Christ has recommended thus far is that you perform certain acts individuated by their outward effects. He starts with omissions: do not resist, do not hit back. He then moves on to positive acts: give what belongs to you, go the extra mile. Again, none of these acts requires that you care about your enemy for their own sake—which is to say, none of these requires love. That revolutionary idea is not introduced until the following passage in the Sermon on the Mount.

Along with being characterized by opposing attitudes, love and hate can be based on opposing reasons.[10] Note that such reasons are not (merely) causal or explanatory; they are considerations *on the basis of which* you might love or hate. That is, you register these considerations as normative, if only implicitly, in such a way that they guide your attitudes.[11]

[10] Strawson (1962: 207) says that attitudes "like resentment, tend to inhibit or at least to limit our goodwill towards the object of these attitudes." More pointedly, *hate* rationally precludes love due to the opposition between the two.
[11] Whether you judge or believe these to be reasons is a further question. You could *believe* that these do not count in favor of love or hate while, for the purposes of those attitudes, nonetheless basing them on such considerations. Think of these attitudes as having their own modules with internal accounting, as it were. Cf. Velleman's (2008) description of the rational relationships between a variety of attitudes like love and hate, which limit when such attitudes can make sense to you.

Taking hate to be merited involves taking there to be reasons for hate sufficient for it to be merited; taking love to be unmerited involves taking there to be reasons against love sufficient for it to fail to be merited.[12] This point can be made fairly ecumenical despite controversy about the nature of reasons for love. First, love and hate need not always be based on reasons. I claim only that, *if* you base love or hate on reasons in a specific instance, those reasons count equally against the opposing attitude for you.

Second, this claim does not rely on a specific view about the reasons for love. Consider two popular, broad views.[13] According to the property view, love is based on features of the beloved that attract you to them.[14] Those who defend the property view like to point out that we are apt to justify love to one another or ourselves by explicitly mentioning certain properties, e.g., "I love her because she's brilliant and has a great sense of humor." It is no less felicitous to invoke repulsive properties as reasons for hate and against love when they are sufficiently egregious: "I hate him for being so selfish and cruel. I could never love him."[15]

Alternatively, or additionally, the relationship view holds that facts about your history with the beloved can serve as reasons for love.[16] Obvious candidates include relations like sharing a past as romantic partners, family members, or close friends. In contrast, your enemy stands in a negative relation toward you, one which counts against love. Whatever relation underlies the enmity in a particular case—e.g., some determinate relation of having hurt your loved ones or resolved to destroy you—can render hate merited and love unmerited. More generally, I suspect

[12] I put this negatively, somewhat tortuously, in order to stay neutral on whether love can be "anti-merited" in addition to "non-merited." See Berker (2024) for an argument against anti-merit.
[13] For an overview friendly to both accounts, see Hurka (2016). Hurka also agrees with my main claim here: reasons of various kinds can count against love as well as in its favor.
[14] See Howard (2019a) for one recent defense among many.
[15] Yao (2020) argues that certain flaws revealing the other's human nature can engender affection and serve as a basis for love. Even if correct, her argument does not plausibly extend to enemies and their most despicable properties.
[16] See Kolodny (2003: 150-151), the best-known proponent of the relationship view: "A's loving B consists (at least) in… believing that A has an instance, r, of a finally valuable type of relationship, R, to person B…"

that any plausible view will share the feature that I have just demonstrated for the property and relationship views: love and hate share the same normative currency.

The past two sections problematize a common refrain: "hate the sin, love the sinner."[17] While I agree that it is appropriate to *dislike* sin, hate as understood here is directed toward a whole person, not a mere action or state.[18] As I argued previously, hate often seems, and sometimes really is, merited toward enemies. This now transmits to love being unmerited if THE HATEFUL ARE UNLOVABLE is correct. Looking at your enemy, it is hard to see how that could fail to be decisive. Still, one might wonder whether all of the reasons to love your enemy are worn on their sleeve. I will return to this question shortly in discussing their human nature.

5. Love Without Reasons

The challenge thus far against LOVE YOUR ENEMY depends on you taking there to be reasons in favor of hate and thereby against love. Let us postpone for now the issue of whether (4) UNLOVABILITY IS DECISIVE. I first want to consider two lines of resistance against what may seem like a rationalistic framing on my part: one is that love for your enemy could be possible if it were irrational, i.e., go against the sum of your reasons (as you see it); a second line of resistance is that love could be arational, i.e., not at all based on (what you take to be) reasons. My reply to both is that such love would fail to be reflective in the cases at issue:

(5) DECISIVENESS PRECLUDES REFLECTIVE LOVE: If you take there to be decisive reason against loving some person *E*, then you cannot reflectively love *E*.

I will consider the prospect of arational love in this section; in the following section, I address irrational love and explain why being reflective is important.

[17] Importantly, this refrain seems to me a misrepresentation of a different thought in Catholic theology; see note 22.
[18] Following Bell (2011), we might call these *globalist* attitudes.

Arational love does not take reasons for or against love into consideration. As before, the proponent of arational love toward your enemy need not claim that all instances of love are arational. They might allow, for instance, that romantic love is based on reasons but maintain that it is possible to love without regard to reasons in other cases. Arguably, the Christian tradition already has a name for arational love: *agape*, the gracious love bestowed on us by God in the first instance.[19] If God is the ethical ideal, a natural inference is that your love for others should look like God's love. What the parable of the Good Samaritan demonstrates, for instance, seems to be a human reflection of *agape* (Luke 10:25-37). The Good Samaritan encounters a stranger in need on the side of the road, sees nothing attractive about him and shares no kinship with him, yet shows love anyways by tending to and sheltering him. To ask what the Good Samaritan found loveable in the stranger is to miss the point. Such love is not based on reasons.

Even if that is the correct interpretation of the Good Samaritan, your enemy is not a ripe candidate for the arational mode of love. After all, your enemy is not a mere stranger to you.[20] Recall that you take them to positively fail to merit love on the basis of hateful features. Ben-Hur registers that Messala is hateful by seeing that Messala has betrayed him, unjustly ruined the Ben-Hur family, and so forth. Once you are in the business of registering reasons for and against loving your enemy, though, arational love vis-à-vis that particular enemy is not available. If you could stop seeing them as unlovable, arational love might then become an option. But in the absence of countervailing reasons or defeaters, the only way to stop seeing your enemy as unlovable is to

[19] It is *agape* that features in the New Testament, not *philia* (paradigmatically, brotherly love) nor *eros* (erotic love). I find it plausible—this is admittedly the arguable part—to categorize *agape* precisely as arational love, i.e., the kind of love not based on reasons. If we further categorize *eros* as love based on intrinsic properties and *philia* as love based on extrinsic properties, we get a neat differentiation of these three central kinds of love that maps onto the three contemporary accounts of reasons for love: the property view, the relationship view, and the no-reasons view. Some philosophers prefer to interpret *agape* as based on something like universal reasons instead, which I discuss later in this section. For more on my suggested taxonomy, see my "Reasons For, Reasons Of, and the Two-Fold End of Love."

[20] A further wrinkle is that animosity existed between Samaritans and Jews in Christ's day, but I abstract away from that aspect of the parable in order to emphasize a contrast between pure strangers and enemies.

delude yourself about what they are like. That is what I mean by unreflective love. By contrast, Good Samaritan or neighbor love need not be unreflective, for you need not ignore anything about a stranger in order to love them arationally. In the next section, I explain why unreflective love is problematic.

In the remainder of this section, it is worth addressing an influential, alternative conception of *agape* as a rational type of universal love.[21] Talk of rational *agape* is itself ambiguous between love of a universal, such as humankind, versus love of a particular based on some universally shared feature. Loving the universal (humankind) is compatible with failing to love or even hating your enemy in particular, which clearly runs afoul of LOVE YOUR ENEMY. It does not help to reply that you love your enemy in particular under a universal description, so long as you still hate them under some more particular description (e.g., loving them *qua* human while hating them *qua* enemy).[22] That state of mind is better described in terms of a more limited pro-attitude (say, admiration) toward the enemy on the basis of one of their features, together with a con-attitude (say, resentment) toward them on the basis of another feature. Our focus, however, is your overall orientation toward the whole person, not privileged under any description: do you love them, or do you hate them? In this sense, love and hate are hegemonic powers in your mental economy. And the hatefulness of your enemy would seem to take priority as far as overall orientation goes.

The proponent of *agape* love toward your enemy might demur, claiming that some universal positive feature really does outweigh hatefulness in all cases. This corresponds to the other possibility for understanding *agape* as rational love: as love of a particular on the basis of

[21] Cherry (2018) advocates that we love perpetrators of racial injustice with this conception apparently in mind.
[22] That seems to be the main thrust of the answer given by Aquinas (ST II-II, q. 25, a. 8). Rather than merely "hate the sin, love the sinner," a better paraphrase of Aquinas would be "hate the sinner, love the human."

universally shared features. Natural candidates include bearing the *imago dei*[23] or being your neighbor in a religious sense;[24] or, in a more secular vein, having a rational nature. Frankly, such thin features strike me as too weak to justify love when weighed against the fact that, say, your enemy has savagely harmed you. Perhaps the point of *agape*, according to this conception, is that it only admits of one consideration in favor of love (the special universal feature) while admitting no considerations against it. But why would that be? It would be more plausible to think that a universal feature could always justify a neutral attitude, like recognition-respect, rather than a positively valenced attitude like love.[25] The latter seems especially implausible when the universal feature—e.g., one's rational nature, the *imago dei*—has been exercised in perverse, awful ways.

David Velleman is well-known for the view that a person's rational nature can be the basis of both respect and love, which he calls the "required minimum and optional maximum responses to one and the same value" (1999: 366). Even if we accept his account, though, it does not entail that love is always available in response to a rational nature. Whereas respect for a person is based solely on the fact that they have a rational nature, that alone does not suffice for love. For the response to reach all the way to love, as Velleman acknowledges, the beloved must manifest their rational nature in particular ways that the lover finds disarming. But that special additional ingredient is present least of all in your relationship with your enemy. Far from being disarming, the enemy's rational nature is manifested in ways that arm you for war (sometimes literally as well as figuratively).

[23] King (1957) says of your enemy, "No matter what he does, you see God's image there. There is an element of goodness that he can never slough off."
[24] See Aquinas (ST II-II, q. 25, a. 6) quoting Augustine.
[25] Neu (2011) makes a similar point in response to Velleman. Indeed, the claim that human nature always calls for *respect* enjoys much wider acceptance than the parallel claim about love.

6. Love Against Reasons

The arguments of the preceding sections entail that you cannot reflectively love your enemy. But why does it matter that love be reflective? In response to the challenge for LOVE YOUR ENEMY, some might urge unreflective forms of love instead. Examining their proposals will help make clear that unreflective love cannot be an ethical requirement, however adventitious it might be. The through-line of this section is a principle in the spirit that "ought" implies "can":[26]

(6) OUGHT IMPLIES REFLECTIVELY CAN: you ought to φ only if you can reflectively φ.

I assume that OUGHT IMPLIES REFLECTIVELY CAN holds for attitudes, love in particular.[27] It is common to qualify the sense of ability at issue in "ought"-implies-"can" as intentional ability, rational ability, or some such.[28] But since loving is not an action, it cannot be done intentionally. Nor do I state the principle in terms of rational ability, so as not to beg the question against arational forms of love such as those just considered. *Reflective* ability is appropriately weaker: it is meant to rule out blind or positively irrational love without requiring that love be based on reasons.[29]

Kierkegaard (1962: 79) might be taken to recommend blind love when he quips:

> Men think that it is impossible for a human being to love his enemies, for enemies are hardly able to endure the sight of one another. Well, then, shut your eyes—and your enemy looks just like your neighbor.[30]

[26] Although assessing this principle relative to a tradition like Christianity is complex, it finds at least some support in verses like 1 Corinthians 10:13, in a passage about moral requirements: "[God] will not let you be tempted beyond what you can bear. But when you are tempted, he will also provide a way out so that you can endure it."
[27] See Portmore (2019) for a relevant discussion of the sort of control over attitudes required for accountability. Note that my assumption is weaker than Portmore's. I only claim that "ought" implies that you *can reflectively* love, regardless of whether or not you *control* your loving.
[28] Cf. Mason (2003).
[29] Gert (2020: 981), while discussing a moral requirement not to seek revenge, endorses an "'ought-implies-rationally-can' principle: the idea that if we are morally required to perform a certain action, it is (at least) rationally permissible (colloquially, not crazy or stupid) to perform it." Given that I take "reflective" to entail "not irrational," the consequent of OUGHT IMPLIES REFLECTIVELY CAN does rule out irrational love. My aim in this section is to offer grounds for this.
[30] Kierkegaard could just be recommending that you ignore features that are in fact irrelevant to whether your enemy merits love (perhaps because all that matters, in his view, is the true underlying nature of being the neighbor). Since I addressed that sort of idea in the previous section, I set it aside in order to focus on another available reading.

An available reading is that Kierkegaard agrees with our earlier conclusion that *agape* as universal love is not, on its own, up to the challenge of loving your enemy. You must first blind yourself to what your enemy is like in order to love them *qua* neighbor, or on the basis of their neighbor status. I have my doubts about the wisdom of this strategy. To extend the metaphor, will blinding yourself not be difficult (and quite painful)? And will you not continue to remember what your enemy looked like from before? Even if you do succeed, your love will not be especially ethical. It will be an expression of a compromised agency by which you are ignorant of how others stand in the world and of your relation to them. That way of living cannot be ethically required of you. Moreover, if you can only love your enemy once you are blind, then the act of blinding yourself in the first place cannot be done out of love for your enemy. The goal is to identify a basis for loving your enemy in the first instance. I take these sorts of concerns to underlie OUGHT IMPLIES REFLECTIVELY CAN, which rules out (this interpretation of) Kierkegaard's suggestion.

A second suggestion for unreflective love comes from Nietzsche (2010: 39), who directs the "noble man" regarding how to think of his enemy:

> To be incapable of taking one's enemies… seriously for very long… Such a man shakes off with a single shrug many vermin that eat deep into others; here alone genuine "love of one's enemies" is possible.

Nietzsche points out that one of the best ways to avoid taking your enemy seriously is simply to forget what they have done against you. But there are numerous problems with this approach as an ethical vindication of LOVE YOUR ENEMY.[31] Forgetting serious wrongs that have been committed against you raises worries about the lack of a proper sense of self-worth. Furthermore, thinking of others as vermin to be shaken off rather than as human beings to be taken seriously is

[31] To be fair, Nietzsche himself likely would not share my interest in an "ethical" vindication of genuine love. Still, the passage is useful for us as a foil.

both mistaken and far from loving. At root, Nietzsche's and Kierkegaard's suggestions share the same basic flaw as *agape* love with respect to the challenge: all of the loves on offer are unreflective, which cannot be ethically required of you.

7. Love for the Wrong Kind of Reasons

To review the bind: If you ought to love your enemy, you must not take the reasons against love to be decisive (according to (5) and (6)). But reflecting on your enemy leads to the conclusion that loving them would be unmerited (according to (2) and (3)). The crux is thus whether merit is decisive, or more precisely:

(4) UNLOVABILITY IS DECISIVE: if you take *E* not to merit love, then you take there to be decisive reason against loving *E*.

Given the preceding discussion, you must look beyond your enemy for reasons that could falsify (4), that is, reasons sufficient to love your enemy. Given the nature of love, however, you cannot love your enemy on the basis of just any reasons. Let me explain why some reasons that really do make it desirable to love your enemy—including prudential, political, and moralistic considerations—are nonetheless the wrong kind of reasons on the basis of which to love.

We risk being misled in this regard by the epigraph from the Sermon on the Mount, where Christ introduces LOVE YOUR ENEMY:

> I tell you, love your enemies and pray for those who persecute you, that you may be children of your Father in heaven… If you love those who love you, what reward will you get?… Be perfect, therefore, as your heavenly Father is perfect. (Matthew 5:44-46, 48)

Apparently, these are various kinds of reasons in favor of loving your enemy. The prospect of a reward is explicitly invoked as a prudential consideration, and the remark about becoming children of God also carries the hint of privileges associated with that status. There may be earthly benefits

in addition, such as esteem from others if you live in a society that honors LOVE YOUR ENEMY. Finally, loving even your enemy may be salutary for your own sense of peace and wellbeing.

To be clear, prudential goods are nothing to sneer at (least of all in the form of rewards from God!). We should agree that they are perfectly legitimate reasons to bring it about that you love your enemy.[32] Various strategies may exist to assist you in this endeavor, ranging from suppressing bad memories of your enemy, to praying for supernatural help. Simply acting as though you love your enemy may eventually lead to actually loving them. Therefore, perhaps the best interpretation of this passage is that Christ is offering proleptic reasons to come to love your enemy—"fake it 'til you make it." What you cannot do directly on the basis of such reasons, however, is *to love*. Love is essentially concerned for the beloved for their own sake. If you were to try to do so for the sake of your own good, you would be loving for the sake of the beloved *for the sake of prudential considerations*, which is incoherent.[33]

Society benefits, too, when you love rather than hate. Martin Luther King Jr. observed this in a commanding sermon about loving your enemy:

> If somebody doesn't have sense enough to turn on the … powerful lights of love in this world, the whole of our civilization will be plunged into the abyss of destruction … Somebody must have sense enough and morality enough to cut off the chain of hate and the chain of evil in the universe. And you do that by love.[34]

[32] My own view is stronger: these are legitimate reasons to love your enemy, albeit not reasons that you can follow directly. But we need not resolve the controversy over whether reasons to φ must be reasons on the basis of which you can φ; see Maguire and Woods (2020) for recent references and discussion. The important point here is that you cannot love on the basis of these reasons.

[33] Cf. Smuts (2014: 519) and Whiting (2017: 404).

[34] King's (1957) sermon goes beyond these points and is a *tour de force* of the many reasons why it is beneficial to love your enemy. Those include the benefits to your own psychic health and the transformative effects of love on your enemy. But even loving your enemy in order to transform them falls prey to the problem I am raising here. On the one hand, you may be transforming them ultimately for the sake of others whom you already care about. On the other hand, if you want to transform them for their own good, then you may indeed be loving them, but the question remains: on what basis?

In this sermon, King explicitly connects his approach of non-violent resistance with LOVE YOUR ENEMY.[35] Consider the social and political gains achieved by the civil rights movement as a result of activists like King resisting hate and striving to love instead. Those are weighty reasons, indeed. But again, on pain of incoherence, they do not give you a basis for caring about your enemy for their own sake. King's rhetoric presupposes that you care about your society but not necessarily your enemy. Adding the presupposition that you care about your enemy will not help, of course, since that is part of precisely what we are seeking reasons for: love.

Christ raises yet another category of reasons, which we might call moralistic reasons, in the passage. In the last verse, he posits the end of being "perfect, therefore, as your heavenly Father is perfect." This is a moralistic reason in the sense that it has to do with morality as such: you have reason to be morally perfect, which requires compliance with everything you ought to do, including LOVE YOUR ENEMY.[36] While moralistic reasons may seem less mercenary than prudential ones, though, they are no more eligible as reasons on the basis of which to love. Again, these are certainly reasons that you can and should endorse, which reflecting on can serve as a basis for bringing yourself into conformity with LOVE YOUR ENEMY. But moralistic considerations cannot be the basis for love itself. Loving the other for their own sake, for the sake of some further end to be brought about, is incoherent no matter how pure the external end.

Here is another way of putting the challenge. Suppose that you act on your prudential, political, or moralistic reasons and succeed in bringing about that you love your enemy. Despite recognizing that they merit hate (assuming you remain reflective), you now sincerely love your

[35] Some of Cherry's other notable work (see especially 2021) should be read with this in mind.

[36] We can also ask why loving your enemy is (part of being) morally perfect. One possibility is that God just so happens to love your enemy, which makes it morally required for you to love your enemy because God sets the moral standard. I prefer instead to think that God is referenced as something like a moral witness, guide or exemplar in this verse. In that case, presumably, there is some further reason in virtue of which loving your enemy is morally desirable for you.

enemy. Our question concerns what you are supposed to be telling yourself from within your newly loving perspective: on what basis does love make sense to you in spite of the reasons that you see against it?

If I am right that the reasons above are the wrong kind of reasons on the basis of which to love, then obviously we should look instead to reasons that are eligible to serve as the basis for love—the so-called "right kind" of reasons.[37] Yet that observation seems to offer little help. Generalizing from above, the wrong kind of reasons count in favor of an attitude in virtue of making it desirable for you to have the attitude; by contrast, considerations that help make an attitude *fitting* are a good starting point for the right kind of reasons.[38] But we now seem to arrive back where we began: your enemy merits hate, not love.

8. Merit Versus Fittingness

To be clear, we seem to be at a dead end because—according to a widespread assumption—only considerations concerning the object of one's attitude can make the attitude fitting and serve as a basis for it.[39] In other words, once your enemy's lack of merit is established, there can be no appropriate basis for loving them. I think that assumption is overly restrictive, however. While I have been arguing that the desirability of an attitude does not help make it fitting, there may be other background facts, also unrelated to the attitude's object, that do help.

For a different attitude, consider grief in a case where a loved one died years ago.

[37] There are several different ways that this point is made in the literature: that the right kind of reasons are *followable*; capable of *guiding* your response; or can serve as a *premise* in your reasoning. See Maguire and Woods (2020) for recent discussion and references.

[38] Cf. Howard (2019b). Note that we need not accept Howard's fittingness-fundamentalism to agree that there is this connection between fittingness, the right kind of reasons, and followability. To be clear, I have not here shown that all fitmakers are necessarily followable, but I proceed on the hunch that this is correct and return to this point later. For other remarks in defense of this connection, see Schroeder (2012).

[39] See Na'aman (2021) for an overview of the "Object View" and how widely accepted it is (though often only implicitly). Na'aman dissents from the object view with regard to grief, which I use as a motivating parallel to love.

Intuitively, it is fitting for your grief to diminish with the passage of time, but not because their death has become less grievous. Rather, it may fittingly diminish because you have already spent years grieving. A pivotal distinction—one that has never to my knowledge been marked explicitly in the literature—must now be introduced. The fact that a certain response is fitting does not, I claim, entail that its object merits the response nor vice versa.[40] I propose that the fittingness orthodoxy is correct regarding *merit* but not fittingness: only merit is determined solely by reasons concerning the object of your response.[41]

This proposal is consistent with the discussion of meriting hate from earlier sections, which turned on features of your enemy. The fittingness relation need not be so constrained, however, for it can depend on facts about the broader circumstances that have nothing to do with the object. The proposal thereby allows for the following plausible combinations of claims: that an object merits, deserves, or is worthy of some response, yet also that the response would not be fitting. The death of your loved one will always merit grief, for example, yet it is no longer fitting to grieve her death given that you have already spent years doing so. Similarly, from the fact that your enemy merits hate and fails to merit love, it does not follow that hate and not love is fitting.

These are admittedly contentious points that I am hoping to motivate. Rather than prove the conceptual distinction in advance, however, I will move forward to demonstrate it in the case of loving your enemy. If I am correct, it will earn its keep. Regarding love in particular, the distinction will allow us to explain how loving your enemy could be fitting in virtue of certain

[40] *Pace* Howard (2023: 2): "The fittingness relation is the relation in which a response stands to its object when the object merits—or is worthy of—that response." I also reject the related claim that a response *x* is fitting only if its object has the corresponding evaluative or aptic property of being *x*-able. Specifically, it is false that it is fitting to love your enemy only if your enemy is lovable.

[41] Note in passing that desert and worth, two other concepts sometimes used interchangeably with fittingness and merit in the literature, group with merit for this distinction: intuitively, an object deserves or is worthy of some response in virtue of its own features. It helps to observe that, grammatically, the natural subject of sentences about merit, desert, and worth is the response's object, whereas the subject of sentences about fittingness is the response itself.

background conditions while maintaining the datum of the previous sections that your enemy does not merit love.

Here is the crucial possibility I have in mind for making it fitting to love your enemy: that you yourself have received unmerited love from a third party at some point in the past. This might be true in a number of ways, ranging from the commonplace that you were cared for as an unpleasant child[42] to the more radical thought that God loves you despite you having positioned yourself as an enemy against God.[43] Controversial though the latter thought may be, it is worth entertaining given its notable influence within Christianity, *inter alia*, where LOVE YOUR ENEMY has deep roots. So suppose for the sake of argument that the following is true:

CRUCIAL ASSUMPTION: you have been loved by another despite meriting hate instead.

The natural thought is that CRUCIAL ASSUMPTION makes it fitting to love your own enemy in turn.[44] Although that could be asserted as a brute principle, I believe that there is more to the story. Something about the similarity between the unmerited love that you received and the unmerited love that you might go on to show your enemy helps to make the latter fitting.[45] Let me now specify two distinct reasons supported by CRUCIAL ASSUMPTION that combine to make love fitting.

9. Welcoming and Exclusionary Reasons

Suppose that you have *welcomed* unmerited love as such. You recognize that it is unmerited and, rather than reject it, you approve of it. I will argue that this gives you reason not to take the

[42] This point is central to feminist care ethics; see, e.g., Kittay (1999). I am unsure, though, whether this alone gets you all the way to LOVE YOUR ENEMY, which might explain why care ethics has given enemy love no special attention.
[43] See Romans 5:8,10: "But God demonstrates his own love for us in this: While we were still sinners, Christ died for us… while we were God's enemies, we were reconciled to him..." Other biblical passages suggest that you can merit God's hate by disregarding God; see, for instance, Romans 1:18-21.
[44] This connection appears persistently in inchoate form throughout the New Testament, as in 1 John 4:19: "We love because he first loved us."
[45] The similarity lies in lack of merit, not necessarily in fittingness. My argument is officially agnostic about whether the original love you received is fitting (it need not be). I address this further in an objection toward the end.

lack of merit of your enemy to be decisive against love, either.

Welcoming love is not the same as experiencing a positive reaction toward some fortunate occurrence like winning the lottery. Enjoying a windfall need not commit you to enjoy the same when it happens to your enemy. When you welcome what another agent *does*, on the other hand, you endorse it in some sense. Hence, we object to welcoming ill-gotten goods such as the gifts of tyrants or donations from morally corrupt organizations. They are unacceptable. By welcoming such things, you are complicit in the giving; you treat the giving as legitimate. Welcoming does not quite commit you to thinking the giving is merited, or even fitting, but to something weaker like permissibility. Thus, you ought not to welcome unmerited good grades, which would undermine the point of grading and unjustly privilege you over others. The same goes for other unmerited goods in social practices that essentially involve justice.[46]

Loving, like literal acts of giving, is not a mere occurrence but rather something that someone *does* (in the broad sense that desiring certain ends, feeling certain ways, and attending to certain things are all what one does, a way of orienting oneself and an expression of one's agency). And, unlike goods such as unfair grades, unmerited love does not essentially undermine justice nor threaten valuable social practices. Thus, it is unobjectionable to welcome unmerited love so long as you do so consistently. But this commits you to being open, in principle, to unmerited love in other cases as well. On pain of inconsistency, you cannot see lack of merit as automatically disqualifying others from love. To hold otherwise would be to make an unjustified exception for yourself. That would be an especially unfitting response to how you were treated.[47]

[46] I am grateful to Jorah Dannenberg for discussion on these points.

[47] The parable of the unforgiving servant in Matthew 18:21-35 speaks to this idea: one servant has a massive debt forgiven yet refuses to forgive the (much smaller) debt owed to him by another servant. In terms of forgiveness, the central message of the parable is that you should forgive others as you have been forgiven. And if you continue to count their hateful features against them, you have not really forgiven. That would be an acutely unfitting response to having been forgiven. Similarly for love, one step in which is forgiveness (see note 50).

CRUCIAL ASSUMPTION alone does not get you this far. After all, you can always refuse to welcome love toward you. The police inspector Javert from *Les Misérables* takes the stance of refusal in dramatic fashion. Having hounded the ex-convict Jean Valjean beyond the bounds of reason, Javert finds the tables turned when the revolution arrives and Valjean has the opportunity to execute him. Javert recognizes that, as enemies, it would be entirely reasonable for Valjean to do so, yet Valjean shows mercy instead. Javert reacts by committing suicide. He does so in part, I think, because he sees that he would otherwise be on the hook to show mercy to others in the same way. Having lived by the law of just deserts, Javert would rather die than accept this result.[48]

Most of us, thankfully, are not like Javert. We tend to welcome mercy and love even when unmerited. Now, that in itself does not yield positive reasons for love. If Javert had accepted Valjean's mercy, he would not thereby be committed never to arrest another fugitive. He would only be committed to seeing mercy as a live option going forward; other cases may turn out to be relevantly different from one's own past. To put this in more contemporary terms, welcoming unmerited love does not necessarily generate first-order reasons for more of the same.

Perhaps CRUCIAL ASSUMPTION functions instead as a *defeater* of first-order reasons against loving your enemy. That is closer to the truth, in my opinion, but still subtly incorrect. The fact that someone has treated you viciously, say, would still count against loving them, just as your own hateful features never ceased being reasons against loving you. These are hard cases of unmerited love precisely because there really is reason against loving your enemy. You can continue to acknowledge that while being committed, *qua* lover, not to hold it against them.

If welcoming love neither gives you a reason to love your enemy nor defeats the reasons

[48] From the musical number "Javert's Suicide": "How can I now allow this man / To hold dominion over me? / This desperate man whom I have hunted / He gave me my life, he gave me freedom / I should have perished by his hand / It was his right / It was my right to die as well / Instead, I live, but live in hell."

against loving them, how does it bear on loving them? In my view, the commitment you gain by welcoming love amounts to an *exclusionary* reason not to take your enemy's hateful qualities into account as reasons against love.[49] This does not mean that you cannot recognize their qualities as hateful, and it certainly does not defeat their status as such.[50]

As a second-order reason, the fact that you have welcomed unmerited love yourself has several noteworthy features. It is the sort of consideration that can bear directly on making up your mind whether to love rather than serving as a first-order reason to act upon yourself. Moreover, it is not a reason to *bring about* anything, which means that it does not face the problem for the wrong kind of reasons mentioned earlier. There, I urged the incoherence of basing a putatively final end on some further final end to be promoted. That you yourself have welcomed unmerited love is a fact about the past, however, not about what to promote. So it is not the wrong kind of reason.[51] Still, this consideration is only a partial basis by which you could love, given that it does not necessarily amount to a positive, first-order reason. You are not yet loving, but you have at least shifted gears and can head in that direction.

10. Gratitude and Expressive Reasons

Positive reason for loving your enemy, I will argue in this section, comes from gratitude. In addition to the exclusionary work above, CRUCIAL ASSUMPTION (a) gives you reason to be

[49] Raz (1990) is the urtext for exclusionary reasons, although there is controversy over the best interpretation of Raz and the most plausible account of exclusionary reasons; see Adams (2021) for a sense of the debate.

[50] This is where forgiveness is relevant. Think, for instance, of parents who forgive their children's killers. The widely publicized case of the Derksens, who announced that they forgave the man who killed their daughter, is discussed by Westlund (2009). While the relationship of forgiveness to love is complicated, I take it that forgiveness involves the forswearing of at least certain attitudes that are in tension with love.

[51] In the earlier section against the wrong kind of reasons for love, I cited Whiting (2017)—but Whiting argues against second-order reasons in general, so why is my proposal here not vulnerable, too? In the passage cited earlier, Whiting (2017: 404) shows only that there cannot be second-order reasons that consist in the desirability of some outcome for love and other non-instrumental attitudes. Whiting fails to distinguish between these and other kinds of second-order reasons, and he mistakenly takes his argument to cut against second-order reasons in general.

grateful, and (b) being grateful, you have reason to extend that love to others. The first step, the reason *for* gratitude in response to the good of unmerited love, must be handled briefly. I take it that being loved is generally good for you. Loving you is a powerful way of *being for* you, which is plausibly valuable for its own sake in addition to any instrumental benefits to you. And receiving what is good provides reason to be grateful to the giver,[52] especially when they are under no obligation.[53] Suppose all of that holds in your case of unmerited love.[54]

The second reason, the reason *from* gratitude, will occupy us longer. Expressions of gratitude to a benefactor can naturally be directed back toward them in more or less fitting ways:

> DIRECT GRATITUDE: if you are grateful to benefactor *B* for a good of type *G*, a fitting expression of your gratitude is to give goods of type *G* back toward *B*.

DIRECT GRATITUDE need not apply to every good turn. If a professional acquaintance gets the check at lunch because it cannot be split, you have reason to pay them back, not necessarily to be grateful; it is perfectly sufficient to pay them back simply for the sake of clearing the debt. Interactions among friends can share this transactional character, too, although usually less formal. Imagine that friends help you move, so you feed them as a way to say thank you.[55] While you can certainly be grateful to them for their help, dinner-for-moving looks more like a moment of communal solidarity where everyone pitches in to make sure that everyone is taken care of ("from each according to their abilities, to each according to their needs"). To be clear, that is appropriate enough and helps avoid an impression of ungratefulness. But it also must be said that neither the

[52] Like welcoming, gratitude in the sense at issue presupposes that there is an agent to whom you are grateful. This is sometimes referred to as triadic or directed gratitude (Manela (2021)): you have gratitude *to* benefactor B *for* good G.
[53] In their survey, Gulliford et al. (2013: 301-2) note that most of the literature on gratitude views the supererogatory nature of the gift as either a necessary condition or an intensifier for gratitude. Later, I address the possibility that loving you was obligatory despite being unmerited.
[54] For a more comprehensive discussion of these and other issues in this section, see my "Love Paid Forward."
[55] I thank an anonymous reviewer for raising this type of scenario and pressing me to clarify why similarity in the type of good matters for expressions of gratitude.

giving nor the response feels especially free; this is not exactly a striking instance of gratitude.

In contrast, suppose a friend graciously buys you a drink, outside the bounds of social scripts and with no expectation of reciprocation; you owe nothing by convention. Here, DIRECT GRATITUDE more clearly gets a grip. What you received was special and heartfelt, and gratitude in response involves appreciation and excitement about that particular kind of good. This accounts for the provision for goods *of the same type* in DIRECT GRATITUDE. Your gratitude for receiving that type of good is what makes it fitting to participate in the giving of the same. Simply paying your benefactor back would suggest that you view the interaction inappropriately, as if it were a transaction. And responding with a different type of good—say, helping your drink-buying friend with moving some furniture—may not be inappropriate, but it fails to express gratitude as fittingly for the special type of good that you received.

You don't always have opportunity to give your benefactor goods of the same type. Perhaps your generous friend passes away unexpectedly after several years of buying you drinks while you were in grad school. You always said how grateful you were, how much you looked forward to buying them drinks once you got a well-paying job, but now they are gone. Nonetheless, you are still able to express your appreciation for that for which you are grateful in the following way:

> EXTENDED GRATITUDE: if you are grateful to benefactor *B* for a good of type *G*, a fitting expression of your gratitude is to give goods of type *G* toward any candidate *C*.[56]

EXTENDED GRATITUDE implies that you have reason to do a good turn to others who are not your departed friend. Intuitively, this is how many people experience the force of gratitude: as an urge to participate in the same type of goodness, whether or not the original benefactor is the new

[56] In fact, I think that EXTENDED GRATITUDE is the more fundamental norm, and DIRECT GRATITUDE is a special instance of it. However, what I argue in the main text does not depend on this point. Regardless of whether EXTENDED GRATITUDE is the first- or second-best option for expressing gratitude, it is the only one applicable in the case of unmerited love. Thanks to Lara Buchak for helpful discussion on the order of explanation here.

recipient.[57] I claim that this sense of moral remainder, so to speak, stands up to reflection as a legitimate reason that makes it fitting to express gratitude.

The same point applies when the benefactor remains in your life but fails to be a candidate for receiving the same type of good. Someone who feels grateful to their mentor in education or vocational training is not typically in a symmetrical situation to do the same for their mentor, so one way they express gratitude is by mentoring others, often in the same type of program.[58] This is compatible, of course, with saying "thank you" directly to the mentor—all of which suggests that EXTENDED GRATITUDE has its own force.

In principle, perhaps DIRECT and EXTENDED GRATITUDE could both apply in response to one and the same good you receive. When it comes to the good of unmerited love, however, DIRECT GRATITUDE is ruled out because someone who gives unmerited love would likely (for that very reason, if for no other) merit love in return. Again, you still can and should love them, but you typically cannot reciprocate the very same good of unmerited love. EXTENDED GRATITUDE, on the other hand, finds a fitting target in your enemy.

What exactly is the first-order reason, then, that makes it fitting to love your enemy? It cannot be that you would be complying with EXTENDED GRATITUDE—that sounds too much like an end to be promoted, which renews worries about the wrong kind of reason. This is how indebtedness works, whether of the literal or moralistic variety: the reason to pay back your creditor is that your debt would then be discharged.

Thankfully, gratitude looks backward rather than forward for its reasons. First, the reason for gratitude is that you received some good. It is not that being grateful will promote a desirable

[57] It is interesting to note that empirical psychological research has repeatedly found adherence to this norm, which is variously labeled "indirect," "generalized," or "upstream reciprocity," as in Grant and Gino (2010). For a very different sort of explanation of upstream reciprocity in terms of evolutionary advantages, see McCullough et al. (2008).
[58] Cf. Kittay (1999: 116) in the care ethics literature.

outcome. Second, expressing some attitude or state does not aim at bringing anything about, yet this does not mean that expressive activity is devoid of reason, either. Expression inherits its reasons from the original reasons *for the attitude*. Compare: expressing anger does not aim at accomplishing anything, but it is not thoughtless. It is motivated by whatever offense was the basis for anger in the first place. Likewise, expressions of gratitude look backward rather than aiming forward. What makes the expressions identified by DIRECT GRATITUDE and EXTENDED GRATITUDE fitting is just whatever supplied the reason *for gratitude* in the first place, namely, the good that you received.

Therefore, CRUCIAL ASSUMPTION can be a reason for which you love your enemy for their own sake because it is a reason not only to be grateful but to express gratitude in addition. Whereas you cannot fully express yourself through DIRECT GRATITUDE when it comes to unmerited love, your enemy is a prime target for EXTENDED GRATITUDE.

Let us take stock. The move to fittingness was prompted by the assumption that reasons on the basis of which you can respond will be those reasons that make the response fitting, albeit not necessarily merited. And intuitively, CRUCIAL ASSUMPTION does help make unmerited love fitting, which called for further explanation. I have identified two distinct types of reasons—exclusionary reasons and reasons of gratitude—that flow out of CRUCIAL ASSUMPTION. And by helping make love fitting, rather than supporting love in virtue of what it would bring about, these bear the tell-tale signs of the right kind of reasons on which to base loving your enemy.[59]

[59] In particular, see Schroeder (2012: 459-461) for further discussion on how right kind of reasons lack the "flavor" of promotion that wrong kind of reasons have, as well as that the right kind of reasons are those that tend to make attitudes fitting. Schroeder agrees, furthermore, that the right kind of reasons are more eligible to serve as bases for your attitudes.

11. Theological Clarifications

Does this solution only work for theological views on which there is a God who loves you first?

This is a solution to the challenge against LOVE YOUR ENEMY only to the degrees that (1) you yourself have failed to merit love and (2) you received love nonetheless.[60] It is unsurprising that these assumptions are both held to be met to a supreme degree in the Christian tradition.[61] That said, religiously inspired though the argument may be, it does not essentially rely on God having been the one who loved you. It works to some degree for anyone who has received unmerited love from anyone else.

Most of us have received unmerited love, and a great deal of love at that, from other humans in our lives. Many of our parents may well have been willing to sacrifice themselves for us just as God does according to Christianity. There are a couple of important caveats, however. First caveat: many of us likely never failed to merit others' love to quite the same degree that we fail to merit God's love according to many traditional religious views. Thus, while both conditions (1) that you yourself have failed to merit love and (2) that you received love nonetheless can be met on a secular version of the solution, I am skeptical that either will be met to the same degree. This may be a theoretical attraction of my view, in fact. Perhaps we should love our enemies, the thought goes, but not our very worst enemies and not to the point of giving up our lives for them, absent the theological scaffolding.

Second caveat: the reason provided by EXTENDED GRATITUDE depends in part on love not

[60] In particular, the reason picked out by EXTENDED GRATITUDE scales to the degree that (1) and (2) are met. Compare McCullough et al. (2008: 281): "Gratitude is responsive to four types of information about the benefit-giving situation: (a) the benefit's costliness to the benefactor, (b) its value to the beneficiary, (c) the intentionality with which it was rendered, and (d) the extent to which it was given even without relational obligations to help (for example, parents' obligations to help their children)."

[61] Ephesians 2:3-5 bluntly captures the extent to which (1) and (2) hold for Christianity: "We were by nature deserving of wrath. But because of his great love for us, God, who is rich in mercy, made us alive with Christ even when we were dead in transgressions—it is by grace you have been saved."

being obligatory for the person who loves you. But if LOVE YOUR ENEMY applied to them, too, then you have less reason to be grateful for what was at least a fitting (if not obligatory) display of unmerited love on their part. In contrast, LOVE YOUR ENEMY will not apply in the same way to God according to my story here (that is, assuming that God has never failed to merit love). If God loves you despite not being called to do so, then you would have more reason to be grateful than you would if you received unmerited love from someone for whom it was nonetheless fitting or obligatory to love you.

If LOVE YOUR ENEMY does not apply to God because God has never been the recipient of unmerited love, that seems to introduce problems for the theological version of the solution. On what basis can God love you reflectively, given an undefeated presumption that there is decisive reason against loving you because you do not merit God's love?

I am not sure how worrisome it would be, were God to love unreflectively. REFLECTIVE LOVE mattered for us in its connection with OUGHT IMPLIES REFLECTIVELY CAN. If it is not the case that God ought to love you, then there is no longer the same pressure for God to be able to love you reflectively. On the one hand, God doing anything unreflectively might not match the conception of God as perfectly rational and omniscient. On the other hand, though, the idea that God loves you unreflectively might fit nicely with a conception of God's love for you as mysterious or even scandalous. While there seem to be a plurality of explanations offered for why you are supposed to love others according to Christianity, God's love for you is not held to the same explanatory demands.

Alternatively, as an explanation for how God could reflectively love you, one set of considerations that arises consistently in religious thought involves the fact that you are God's

creature, or that you stand in a number of metaphorical relationships to God (such as being God's child). That may furnish a special explanation for how God could reflectively love you even if you do not merit it. And because it is generally not the case that your enemy stands in any of those relationships to you (assuming that they are neither your child nor your creation), such reasons for God to love you would not apply to you vis-à-vis your enemy. Thus, the best strategies for making sense of God's love for you and your love for your enemy would diverge.

12. Conclusion

In conclusion, let me reiterate payoffs on several levels. One is practical (in the loose sense that guidance for attitudes is practical): insofar as the solutions that I have outlined really are reasons that make it fitting to love your enemy, drawing your attention to them enables you to reflect further and, just maybe, actually love the enemies you encounter in life on those bases.

Another payoff is theological. LOVE YOUR ENEMY is prominent in religious traditions such as Christianity, and the solution that I suggest on its behalf also turns out to be well-grounded in Christianity. So that central plank in a certain type of religious ethical framework is defensible.

Lastly, there is an important metaethical payoff. If my proposed solutions are correct, their success bears on the scope of the right kind of reasons for attitudes more generally. Philosophers widely hold that the right kind of reasons must directly depend only on considerations involving the object of the attitude. Yet there exist reasons of the right kind to love your enemy, I argue, that depend on relationships with others apart from your enemy. If this is correct, then there can be reasons of the right kind that do not depend on the attitude's object—and that do not make the attitude merited—but that nonetheless help make it fitting in virtue of the broader circumstances.

The reasons that I identified for loving your enemy are based on CRUCIAL ASSUMPTION: that you have received unmerited love yourself. That assumption is the source for two distinct

reasons on the basis of which you can love your enemy, which I presented as a cumulative case. First, insofar as you have welcomed unmerited love, you find yourself committed not to count your enemy's lack of merit as decisively disqualifying them from love. This is a higher-order reason to exclude first-order reasons against loving your enemy from consideration. Second, if you have received unmerited love, you have reason not only to be grateful but to express your gratitude by extending unmerited love to others. Who better to start with than your enemy?[62]

[62] I am very grateful for feedback on prior versions of this paper from Christopher Bobonich, Michael Bratman, Lara Buchak, Andrew Chignell, Jorah Dannenberg, Nadeem Hussain, Barry Maguire, Anthony McDougal, Z Quanbeck, Leif Wenar, and a pair of anonymous reviewers for OSPR. It has also greatly profited from questions and comments from audiences at Princeton University's Center for Human Values, the Emotion and Society Works-In-Progress Series, the Society of Christian Philosophers, and the Theistic Ethics Workshop.

References

Antill, Gregory. 2020. “Epistemic Freedom Revisited.” *Synthese* 197 (2): 793–815.

Adams, N. P. 2021. “In Defense of Exclusionary Reasons.” *Philosophical Studies* 178 (1): 235–53.

Aquinas, Thomas. 1920. *Summa Theologiæ*. Originally published 1265–1274. Translated by Fathers of the English Dominican Province. Edited by Kevin Knight. New Advent. Accessed March 3, 2025. https://www.newadvent.org/summa/.

Bell, Macalester. 2011. “Globalist Attitudes and the Fittingness Objection.” *The Philosophical Quarterly* 61 (244): 449–72.

Berker, Selim. 2024. “Is There Anti-Fittingness?” *Ergo* 11 (39): 1051–82.

Cherry, Myisha. 2018. “Love, Anger, and Racial Injustice.” In *The Routledge Handbook of Love in Philosophy*, edited by Adrienne M. Martin: 157-168. Routledge.

———. 2021. *The Case for Rage: Why Anger Is Essential to Anti-Racist Struggle*. Oxford University Press.

D’Arms, Justin, and Daniel Jacobson. 2000. “The Moralistic Fallacy: On the ‘Appropriateness’ of Emotions.” *Philosophy and Phenomenological Research* 61 (1): 65–90.

Ebels-Duggan, Kyla. 2008. “Against Beneficence: A Normative Account of Love.” *Ethics* 119 (1): 142–70.

Frankfurt, Harry G. 2004. *The Reasons of Love*. Princeton University Press.

Gert, Joshua. 2020. “Revenge Is Sweet.” *Philosophical Studies* 177 (4): 971–86.

Grant, Adam M., and Francesca Gino. 2010. “A Little Thanks Goes a Long Way: Explaining Why Gratitude Expressions Motivate Prosocial Behavior.” *Journal of Personality and Social Psychology* 98 (6): 946–55.

Gulliford, Liz, Blaire Morgan, and Kristján Kristjánsson. 2013. "Recent Work on the Concept of Gratitude in Philosophy and Psychology." *The Journal of Value Inquiry* 47 (3): 285–317.

Howard, Christopher. 2019a. "Fitting Love and Reasons for Loving." In *Oxford Studies in Normative Ethics Volume 9*, edited by Mark Timmons: 116-137.

———. 2019b. "The Fundamentality of Fit." In *Oxford Studies in Metaethics Volume 14*, edited by Russ Shafer-Landau: 216–36.

———. 2023. "Forever Fitting Feelings." *Philosophy and Phenomenological Research* 107 (1): 80-98.

Hurka, Thomas. 2016. "Love and Reasons: The Many Relationships." In Love, Reason and Morality, edited by Katrien Schaubroeck and Esther Kroeker: 163–80. Routledge.

Kierkegaard, Søren. 1962. *Works of Love*. Translated by Howard V. Hong and Edna H. Hong. Harper and Brothers.

Kittay, Eva Feder. 1999. *Love's Labor: Essays on Women, Equality and Dependency*. Routledge.

Kolodny, Niko. 2003. "Love as Valuing a Relationship." *The Philosophical Review* 112 (2): 135–89.

Lewis, C. S. 1960. *The Four Loves*. Harcourt, Brace.

Maguire, Barry, and Jack Woods. 2020. "The Game of Belief." *The Philosophical Review* 129 (2): 211–49.

Manela, Tony. 2021. "Gratitude." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2021/entries/gratitude/.

Marxen, Chad. 2021. "Closing the Case on Self-Fulfilling Beliefs." *Australasian Journal of Philosophy* 101 (1): 1–14.

Martin Luther King, Jr. 1957. "'Loving Your Enemies,' Sermon Delivered at Dexter Avenue Baptist Church." The Martin Luther King, Jr., Research and Education Institute. February 11, 2015. https://kinginstitute.stanford.edu/king-papers/documents/loving-your-enemies-sermon-delivered-dexter-avenue-baptist-church.

Mason, Elinor. 2003. "Consequentialism and the 'Ought Implies Can' Principle." *American Philosophical Quarterly* 40 (4): 319–31.

McCullough, Michael E., Marcia B. Kimeldorf, and Adam D. Cohen. 2008. "An Adaptation for Altruism: The Social Causes, Social Effects, and Social Evolution of Gratitude." *Current Directions in Psychological Science* 17 (4): 281–85.

McDougal, Austen. Manuscript. "Love Paid Forward."

———. Manuscript. "Reasons For, Reasons Of, and the Two-Fold End of Love."

Na'aman, Oded. 2021. "The Rationality of Emotional Change: Toward a Process View." *Noûs* 55 (2): 245–69.

Neu, Jerome. 2011. "On Loving Our Enemies." In *The Ethics of Forgiveness*, edited by Christel Fricke: 130-142. Routledge.

Nietzsche, Friedrich. 1989. *On the* Genealogy *of Morals and Ecce Homo*. Originally published 1887. Translated by Walter Kaufmann. Vintage.

Portmore, Douglas W. 2019. "Control, Attitudes, and Accountability." In *Oxford Studies in Agency and Responsibility Volume 6*, edited by David Shoemaker: 7–32. Oxford University Press.

Raz, Joseph. 1990. *Practical Reason and Norms*. Oxford University Press.

Schönberg, Claude-Michel, Alain Boublil, and Herbert Kretzmer. 1985. *Les Misérables: The Complete Libretto*.

Schroeder, Mark. 2012. "The Ubiquity of State-Given Reasons." *Ethics* 122 (3): 457–88.

Schwab, George. 1987. "Enemy or Foe: A Conflict of Modern Politics." *Telos* 1987 (72): 194–201.

Smuts, Aaron. 2014. "Normative Reasons for Love, Part II." *Philosophy Compass* 9 (8): 518–26.

Strawson, Peter Frederick. 1962. "Freedom and Resentment." Proceedings of the British Academy 48: 187–211.

Velleman, J. David. 1999. "Love as a Moral Emotion." *Ethics* 109 (2): 338–74.

———. 2008. "A Theory of Value." *Ethics* 118 (3): 410–36.

Westlund, Andrea C. 2009. "Anger, Faith, and Forgiveness." *The Monist: An International Quarterly Journal of General Philosophical Inquiry* 92 (4): 507–36.

Whiting, Daniel. 2017. "Against Second-Order Reasons." *Noûs* 51 (2): 398–420.

Wyler, William, dir. 1959. *Ben-Hur*. Metro-Goldwyn-Mayer.

Yao, Vida. 2020. "Grace and Alienation." *Philosophers' Imprint* 20 (16): 1–18.