THE UNIVERSITY OF CHICAGO

IN PRAISE OF PRAISE

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE HUMANITIES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHILOSOPHY

BY

DANIEL J. TELECH

CHICAGO, ILLINOIS

AUGUST 2018

ACKNOWLEDGMENTS

I thank my dissertation committee. I am incredibly fortunate to have had Agnes Callard and Brian Leiter direct my dissertation project. Their support, wisdom, and generosity have meant a great deal to me—philosophically and personally—over the past several years. Joint dissertation meetings with Agnes and Brian unfailingly left me with a sense of urgency, demandingness, and encouragement that remains with me, at least on good days, when doing philosophy. Before they were my advisors, they were my teachers. Agnes' seminar on deliberation, on the one hand, and Brian's workshop on free will and responsibility, on the other, played significant roles in my becoming gripped by the questions animating this project. I hope to be able to live up to the ideals that working and studying with them has allowed me, however incipiently, to appreciate.

I thank Paul Russell for being an excellent committee member. Paul has helped me stay attuned to the complexity and humanness of issues of agency and responsibility. This dissertation owes much to insightful conversations with him. I also thank Derk Pereboom. Derk supervised a valuable visit of mine to Cornell in the fall of 2016, and became something of an unofficial committee member, providing me with generous and instructive comments on the majority of the dissertation.

There are many others to whom I am grateful for support with and valuable discussion on, parts of my project, and its earlier and attendant stages. I thank Nomy Arpaly, Dani Attas, Gunnar Björnsson, Dan Baras, Alex Boeglin, Matt Boyle, Pascal Brixel, Mark Bronson, Dan Brudney, Andreas Carlsson-Brekke, Justin Coates, Taylor Cyr, Evan Dutmer, David Enoch, Augie Faller, Paul Franks, Tobias Fuchs, Sol Goldberg,

ABSTRACT

It is a commonplace that blameworthy agents deserve blame, and praiseworthy agents deserve praise. But while blame has in recent years received considerable and often illuminating philosophical discussion (e.g. Bell 2013; Coates and Tognazzini 2013; Hieronymi 2004; Mason 2011; McKenna 2012; Pickard 2014; Scanlon 2008; Sher 2008; Smith 2007; Talbert 2012; Wallace 1994, 2011; Wolf 2011), the nature and norms of praise remain relatively unanalyzed. Since the publication of Peter Strawson's highly influential essay "Freedom and Resentment" (1962), philosophers have increasingly understood moral responsibility in terms of our susceptibility to a range of moral emotions (or "reactive attitudes") in response to expressions of good and ill will. The Strawsonian, moral psychological, approach to moral responsibility owes part of its appeal to its ability to captures the valence and interpersonal significance of the forms of assessment characteristic of our responsibility responses. But while Strawson himself includes blaming and praising emotions among the reactive attitudes, subsequent philosophical work in this area has concentrated overwhelmingly on blame. Our philosophical vocabulary of praise is impoverished. (But this is not because we fail to respond positively to praiseworthy actions. Nor do we fail to recognize the *importance* of so responding. The popular saying, 'give credit where credit is due' is an injunction to just this.) This omission matters in part because blameworthiness and praiseworthiness are opposed modes of moral

responsibility—i.e. they are taken to be two—often *the two*— ways of being morally responsible.[1]

A negative effect of the disproportionate attention dedicated to blame is that philosophers have imported to their conceptions of moral responsibility *in general* features that are in fact specific to blame(-worthiness). Resultantly, we not only lack an account of the nature and norms of praise; philosophical work on our responsibility practices displays a narrow, quasi-legalistic, moral orientation, one in which the concepts of normative demand, right-violation, and sanction reign supreme. If, however, we seek to understand what it is to be, and to respond to, a full range of responsible agents, it is incumbent upon us to enrich our vocabulary of praise. My dissertation takes a step in this direction.

In chapter 1, I offer a novel account of gratitude and on its basis argue that the significance of benevolence is limited in several ways that have not been duly appreciated by pure quality of will theorists of moral responsibility—i.e. those who think that quality of will is *all* that matters in assessing an agent's moral responsibility—and especially those quality of will theorists who take the reactive attitudes to comprise our responsibility

---

[1] This is not a claim merely about recent scholarship about moral responsibility. The idea that expressions of our agency may be assessable as warranting blame or praise—where these are understood as contraries— has featured in philosophical thought concerning agency and ethics since at least the 4th century BC. In the *Eudemian Ethics*, Aristotle (2.1223a) writes that "since goodness and badness and the actions that spring from them are in some cases praiseworthy and in other cases blameworthy (for praise and blame are not given to things that we possess from necessity or fortune or nature but to things of which we ourselves are the cause, since for things of which another person is the cause, that person has the blame and the praise), it is clear that both goodness and badness have to do with things where a man is himself the cause and origin of his actions." Epicurus similarly writes: "...some things happen of necessity, others by chance, others through our own agency...necessity destroys responsibility and chance is uncertain; whereas our own actions are autonomous, and it is to them that praise and blame naturally attach." (*Letter to Menoeceus*, §133)

responses, i.e. Strawsonians. The claim that I defend most extensively in that chapter is that gratitude is susceptible to retroactive feedback, according to which the (unforeseen and unintended) consequences of the benefactor's manifestation of quality of will can give the beneficiary reason to be *more grateful* to their benefactor than they were for the immediately willed benefit, despite the fact that the benefactor's degree of benevolence is recognized as remaining diachronically stable. If this "gratitude-feedback thesis" is accepted, in those cases where gratitude constitutes a form of moral praise, as Strawsonians maintain it sometimes does, it will be a matter of luck in consequences whether an agent's action garners her more or less praise. Another way to put this is to say that the moral worth of an action, its praiseworthiness, may be vulnerable to resultant moral luck (Nagel 1979; Williams 1981).

In chapter 2, I provide an analysis of the nature of pride that overcomes the shortcomings of the *restrictive approach* to pride, on the one hand, and the *disjunctive approach*, on the other. On the restrictive approach, pride involves viewing an object as expressive of one's agency and as something for which one is thereby *morally responsible*. On this view, to be proud of one's naturally good memory is either to misrepresent one's memory as a product of one's agency, or to misidentify one's *happiness that* one has a good memory. On the disjunctive view, by contrast, there are really two distinct emotions called "pride": a moral emotion and a non-moral one. Although the disjunctive view accommodates pride for traits and other non-actions, it fails to explain the sense we have that there is unity to our range of pride emotions. A capacious and unified view of pride's objects is available, and through its outline and defense I shed light on the relationship

between agency and practical identity, one reflecting that we are, even as agents, more than what we do.

In chapter 3 I argue that interpersonal praise cannot be understood in terms of deontic concepts like "moral demand", and that the assumption that it can be so understood—which I call the *demand prejudice*— has a stronghold largely owing to the kind of dominance that "accountability" views of responsibility have enjoyed. I propose a novel axiological conception of responsibility in the *estimability* sense. Estimability characterizes our aspirational stance of holding agents *up* to interpersonal ideals. It is in light of this latter stance, which I characterize as a broad disposition of "normative hope", that our praise of others is fruitfully understood. I then provide a way of unifying accountability and estimability by construing both as forms of responsibility in the *deserved moral address* sense. On this picture, accountability blame and estimability praise are both paradigmatically expressed through moral address (demand and affirmation, respectively), which forms of address characteristically impact their targets' interests.

In chapter 4, I challenge the assumption that blame is unique in being governed by norms of "standing", like the non-hypocrisy condition. I first motivate a shift in attention from the target of praise to the person doing the praising, i.e. the praiser, by introducing two relatively uncontroversial norms governing the praise of candidate praisers: an epistemic condition and a right-relationship condition. I then make a case for the existence and importance of norms of standing to praise. Specifically, I argue that the praise of candidate praisers is governed by what I call the "familiarity condition": *for S to appropriately praise T for ϕ-ing, S's praise must be based in familiarity with the value expressed in T's ϕ-ing*. Praise that fails to satisfy this condition is vulnerable to a distinctive

kind of criticism, one based in considerations *comparable to*, but importantly distinct from, those underlying the inappropriateness of hypocritical blame. Reflection on the familiarity condition sheds light on a valuable feature of our responsibility practices that is otherwise overlooked, namely the phenomenon of co-valuation that successful praise engenders.

As a whole, *In Praise of Praise* proposes that we think of moral responsibility as bearing not only on questions of how, who, and under what conditions we are justified to blame, but on questions about who we— especially as participants of a range of ideal-governed relationships— aspire to be.

CONTENTS

1

Gratitude, Praise, and Moral Luck

0. Introduction

Gratitude is a response not simply to beneficence, but also to bene*volence*. A difference of moral significance— philosophers and non-philosophers agree[1]— is made by the quality of the benefactor's motivations. Natural events and malicious persons may both be causally responsible for the fulfillment of some end of mine (i.e. they may *benefit* me), and while I may resultantly be appreciative *that* the beneficial state of affairs came about, I am ordinarily not thereby disposed to feel grateful *to anyone*. This is plausibly because I do not take these benefits to be products of benevolence, or good will. The claim that gratitude includes viewing another to have acted from good will is a mainstay in philosophical discussion of gratitude (e.g. Berger 1975; Camenisch 1981; Roberts 2003: 294-295, 2015). Additionally, on the increasingly prevalent Strawsonian approach to moral responsibility, it is in virtue of gratitude's responsiveness to others' quality of will that it belongs to a class

---

[1] And of the non-philosophers, "for both adults and children, malicious intent disqualified gratitude." See Gulliford and Morgan (2014: 20).

of emotions that bear an essential relation to what it is to be, and to hold others, morally responsible.

Strawson's thought is this: it is an ineliminable feature of human social life that we care about the regard or good will (and disregard, which includes but is broader than, ill will) that others show us, and our so caring is registered in the disposition to feel and express certain emotions in response to others' regard or disregard. These emotions, which Strawson labels "reactive attitudes", "rest on, and reflect, an expectation of, and demand for, the manifestation of a certain degree of goodwill or regard on the part of other human beings" (Strawson 1962: 84). The negative reactive attitudes (resentment, indignation, and guilt) have the affective and motivational profiles we associate the blame, and the positive reactive attitudes (gratitude, admiration, and pride) have the affective and motivational profiles of praise. In contrast to theorists who take praise and blame to be a matter of *judging* blameworthy or praiseworthy, the Strawsonian is able to capture the interpersonal significance, and valence, of the forms of appraisal involved in our responsibility responses. For, the reactive attitudes are "forms of emotional reaction that mark the moral meaning of others' morally significant actions" (Macnamara 2011: 89); in virtue of this they are "vehicles whereby we hold morally responsible" (McKenna 2012: 2). Since these attitudes have their basis in a disposition that is practically and rationally ineliminable, and as such stands in no need of external justification, our practices of praise and blame are immune to metaphysical threats, most notably the threat determinism. The truth of determinism has no bearing on whether a person manifests good or ill will, and it is to considerations about agents' quality of will that emotions like gratitude and resentment are sensitive.

Is quality of will, however, the *only* factor that reactive attitudes are properly sensitive to? An affirmative answer to this question is arguably implicit in Strawson, and contemporary Strawsonians have advanced the view that our responsibility responses (i.e. attitudes of praise and blame) are appropriately responsive only to quality of will, e.g. benevolence, malevolence, indifference, and the like. That is, they have tried to defend a Strawsonian *pure quality of will view*.[2] Quality of will can be unpacked in a variety of ways, and we need to settle on a particular understanding here. The idea that the agent's quality of will is the sole determinant of an action's moral worth is often associated with Kant (GM 4:394). But the Kantian, who identifies good quality of will with the 'motive of duty', is as much within my purview as is the philosopher who identifies good quality of will with intrinsic desire for what is good or right (Arpaly & Schroeder 2014). 'Will' is a capacious term put to use by proponents of various views of moral motivation. What matters here is

---

[2] See Shoemaker (2011; 2015), but I take this to be the default view of those who hold Strawsonians view of responsibility, especially ones with Kantian commitments (Darwall 2005; Wallace 1994). A pure quality of will view is arguably held by McKenna 2012 (though see Shoemaker 2011: 99 n.14); Rosen (2015); Macnamara 2011; 2013). I take no stand here on whether Strawson is best understood as a pure quality of will theorist. Certainly, he emphasizes that "[t]he reactive attitudes…are *essentially* reactions the quality of other's will toward us, as manifested in their behavior: to their good or ill will or indifference or lack of concern" (italics added). Furthermore, Strawson holds that when agents are excused and exempt from attitudes of blame or praise, there being so excused or exempt is explained by considerations about quality of will, specifically considerations revealing that—owing to factors like ignorance, incapacity, or cognitive deficiency— the agent's action lacks the relevant basis in his will for reactive attitudes (Strawson 1962: 78-9). In other words, quality of will seems to be, at least put forward as, doing all the justificatory work (– though see Watson (1987) for an influential worry). Though not himself a Strawsonian, Zimmerman (2015: 51) attributes to Strawson a pure quality of will view, writing that for Strawson there is "a single answer" to the question 'what grounds ascriptions of moral responsibility," namely "the person's *quality of will* as manifested in his behavior." Of compatibilists who are generally sympathetic to a Strawsonian approach, Russell (2017a) is an exception in explicitly rejecting a pure quality of will view. I return to this near this paper's end.

the pure quality of will theorist's contention that our praise and blame should be determined only by the features of the agent's moral motivational psychology (*whatever this is*), as manifested in his action. Since I will use 'good will' (/'good quality of will') and benevolence interchangeably, benevolence too should be understood in this noncommittal sense.

The aim of this paper is to assess the pure quality of will view by attending to the nature of gratitude. I offer a novel account of gratitude and on its basis argue that the significance of benevolence is limited in three ways that have not been duly appreciated by pure quality of will theorists of moral responsibility (henceforth, simply, *quality will theorists*)—i.e. those who think that quality of will is *all* that matters in assessing an agent's moral responsibility—and especially those quality of will theorists who take the reactive attitudes to comprise our responsibility responses, i.e. Strawsonians.[3] In particular, I argue that i) the import of benevolence is constrained by the value of the act-type the benefactor seeks to bring about, and ii) the sensitivity of praise to *effort undergone* limits the justificatory role benevolence can occupy by rendering gratitude and praise vulnerable to "circumstantial luck" (Nagel 1979: 28). The claim that I most extensively defend, however, is that iii) gratitude is susceptible to retroactive feedback, according to which the (unforeseen and unintended) consequences of the benefactor's manifestation of quality of will can give the beneficiary reason to be *more grateful* to their benefactor than they were for the immediately willed benefit, despite the fact that the benefactor's degree of benevolence is recognized as remaining diachronically stable. If this "gratitude-feedback

---

[3] The first two limitations apply equally to all quality of will theorists (i.e., regardless of whether they are Strawsonians).

thesis" is accepted, in those cases where gratitude constitutes a form of moral praise, as Strawsonians maintain it sometimes does, it is a matter of luck in consequences whether an agent's action garners her more or less praise. Another way to put this is to say that the moral worth of an agent's action, its praiseworthiness, may be vulnerable to resultant moral luck (Nagel 1979; Williams 1981). Contra the quality of will theorist—according to whom an agent's moral responsibility (praiseworthiness and blameworthiness) is determined *solely* by the quality of will manifested in the agent's action——I argue that gratitude-feedback cannot be discounted as an irrational feature of our moral psychologies. The quality of will theorist's commitment to the principle that we can only be praised for what is expressive of our good will should therefore be rejected. By directing our attention to cases of *good* moral luck— which I argue are invulnerable to a key objection levied against their blame-involving counterparts— I create dialectical trouble for the wholesale opponent of moral luck.

I proceed as follows. After developing independently plausible moral psychological account of gratitude—one that does justice to the quality of will theorist's intuition that benevolence is an important feature of gratitude's objects (and so, at least, *a* determinant of moral worth) (1), I argue that our moral lives are expressive of the fact that we recognize bounds on the value of benevolence (2).

Before all this, however, let me clarify why I focus here on *gratitude*. One reason is that work on moral responsibility, the Strawsonian literature notwithstanding, tends to focus on attitudes of blame. Our understanding of and intuitions about holding responsible are for this reason apt to be skewed in the direction of blame. Given that Strawson takes gratitude and the paradigmatic attitude of blame— *resentment*— to be a "usefully opposed

pair" (Strawson 1962: 77), Strawson's landmark essay might have titled, "Freedom and Gratitude". Furthermore, since praise and blame are sometimes thought to have different control conditions (Wolf 1980, 1990; Nelkin 2008, 2009; Vilhauer 2015), a shift in focus to praise-manifesting attitudes may afford a novel perspective on the importance of quality of will in holding responsible. Lastly, quality of will theorists working on "moral worth" (Arpaly 2002; 2003; Markovits 2010; Sliwa 2016) too propose that worthiness of praise be understood in terms of the agent's motives. By attending to a praise-manifesting attitudes, we avail ourselves of these interlocutors.

In moving back and forth between talk of 'gratitude' and talk of 'praise', however, I am neither identifying the two nor presupposing that being grateful is always way of praising or giving credit. Quality of will theorists who are not taken with Strawson (e.g. Arpaly 2002; 2003; 2006) will resist the idea of making inferences about the nature of praise on the basis of the nature of gratitude. They may point to examples in which it is reasonable to be grateful to someone, but not to praise them (or even judge them praiseworthy).[4] That's fine. I am here interested in those cases of gratitude that are expressive of praise (or at least *track praiseworthiness*), and so, might be called instances of *moral gratitude*. They are instances of being grateful to another for benefitting one from morally good will. The fact that our reactive attitudes are not always or ideally sensitive to quality of will leaves intact the idea that such attitudes may be vehicles of praise and blame, or at least, *evidence* of praiseworthiness or blameworthiness. (Especially if we grant that

---

[4] "We might feel more gratitude toward the person who washed the dishes today for the first time than we ever feel toward the person who does it every day, whom we take for granted… if we do not happen to have the time and inclination to reflect on our judgments, excessive gratitude easily turns into an overestimation of the dishwasher's praiseworthiness" (Arpaly 2006: 29).

we are also fallible in our judgments about blame and praise.) In any case, the scope of the following discussion is not limited to those who take gratitude to be a vehicle of praise (e.g. Strawson (1962); Darwall (2006); McKenna (2012); Macnamara (2011; 2013); Shoemaker (2011; 2015); Rosen (2004: 311); Smith (2011); Martin (2013); Russell (2013)); it includes non-Strawsonians who will grant that "moral emotions are an important source of *intuitions about* praiseworthiness and blameworthiness" or that "[t]he reactive attitudes are readily felt *signs* of what is morally significant" (Arpaly 2006: 29; 31, italics added).[5]

1. The Marks of Gratitude

The aim of this section is to provide an account of the moral psychology of gratitude. I am not directly concerned with the normative question of when it is *appropriate* to have the attitude—in the sense that it would typically be *inappropriate* to be grateful to the doctor for wearing a cardigan or for poisoning one's best friend—but what might be called its intelligibility-conditions: what must be true of some person for us to intelligibly attribute to him the emotion, gratitude? I identify three necessary conditions: i) the benevolence condition; ii) the appreciation condition; iii) the expressive condition. While I think these conditions are jointly sufficient for gratitude, all that matters for my purposes is that we take them to be necessary conditions. And since they are necessary conditions that are

---

[5] In the literature on moral worth we are regularly brought to reflect on the admirableness of agents and actions, as well as whether we'd be grateful to then, in forming judgments about their moral worth. Indeed "admirable" is sometimes used in ways that suggest it is meant to be interchangeable with "praiseworthy" (see esp. Markovits 2012: 289-290; 294; 297; 299; 301n. 26; 303-4; 207; 310).

revelatory of gratitude's nature (unlike the condition that 1+1= 2, which too will necessarily be satisfied in any world in which some agent counts as grateful), I will refer to the features they describe as the essential features, or "marks" of gratitude.

Before launching into discussion of the marks of gratitude, I note that there is a usage of the word 'gratitude' that does not pick out an inherently interpersonal, or social emotion. This is the usage according to which one might say, 'I am grateful *that* I got to see a shooting star.' This sense of gratitude, sometimes called 'propositional gratitude', involves a relation between a person and a proposition (McAleer 2012). This is not my topic. By "gratitude" I mean the social and essentially agent-directed emotion that involves a triadic relation between two agents and (typically) an action, of the following sort: "Abe is grateful to Miranda for helping him move into his new apartment". Regardless of whether 'propositional gratitude' is gratitude only in name—little seems to be lost by redescribing instances of 'propositional gratitude' as cases of *appreciation* [6] —by "gratitude" I mean the interpersonal attitude that is essentially agent-directed.[7]

1.1. *The Benevolence Condition*

A mark of S's being grateful to T for ɸ-ing is satisfaction of

---

[6] Cf. Carr (2013), Roberts (2015) and Manela (2016*)*.

[7] Some statements of (*so called*) propositional gratitude are elliptical for statements of prepositional gratitude. A brief anecdote: walking to Harper Reading Room one afternoon, a couple locked in arms walked by, and the women said, "I'm so grateful that you even thought of me for that." I don't suggest that all personally-directed statements of propositional gratitude can be re-written as statements of prepositional gratitude, but I suspect that my passerby's statement is translatable into the following: "I am so grateful to you for having thought of me for that." Of course, this needn't be true of my actual passerby for the point to stand.

The Benevolence Condition, i.e. that S takes T to have ɸ-d from good will/ benevolence.[8]

Strawson appears to accept the benevolence condition when he writes:

> If someone's actions help me to some benefit I desire, then I am benefited in any case; but if he intended them so as to benefit me because of his general good will towards me, I shall reasonably feel gratitude which I should not feel at all if the benefit was an incidental consequence unintended or even regretted by him, or some plan or action of a different aim. (Strawson 1962: 76)

The benevolence condition is accepted not only by Strawsonians, but by philosophers of emotion who are not engaged in the debate over moral responsibility (e.g. Berger 1975; Camenisch 1981; Roberts 2003: 294-295, 2015). Benefits certainly seem to have a special significance when they are *meant* by another to benefit us. Consider Alex, a poor swimmer whose life is in the grip of violent waves. Upon being rescued, Alex is the recipient of beneficence regardless of whether her benefactor (Ben) saves her out of i) malice, ii) indifference, or iii) concern for her well-being. Alex is benefitted regardless of Ben's motives, but if Ben saves Alex from concern for her well-being, Alex is the recipient not only of beneficence, but also benevolence. Her benefactor acts in response to moral reasons—and for this reason Ben would, it seems, be the fitting target of Alex's gratitude.

A few qualifications about the benevolence condition are in order. First, while in many cases the subject who feels gratitude is the target of benevolence, this need not be so. Sometimes we feel grateful for the good will directed toward those with whom we identify. Consider for instance the father's gratitude to his son's dedicated teacher. The father, we can stipulate, knows that the teacher acts out of good will towards her pupil, and

---

[8] —where S and T stand for distinct agents and ɸ stands for some act token. e.g., 'Sue (S) takes Tom's (T's) restarting Sue's stalled car (ɸ) to have been done from benevolence towards her.'

not towards him.[9] While it is possible that the father construes himself as the ultimate beneficiary of the teacher's goodwill, there is little reason to think that he must do so in order to feel grateful to the teacher.

Second, While I refer to the benevolent attitude and the benefit (the φ-ing) as normatively distinct phenomena, one might think that benevolent attitudes can *themselves* constitute benefits, and so, sometimes be the proper objects of gratitude. Indeed, if one thinks that good will is the *only thing* that has moral significance one might think that the 'benevolent attitude' and 'benefit' share the same intension. This claim has roots in Stoic thought and is endorsed by Seneca, who writes: "[a] benefit cannot be touched with one's hand; the business is carried out with one's mind. There is a big difference between the raw material of a benefit and the benefit itself… Consequently, the benefit is not the gold, the silver, or any of the things which are thought to be most important; rather, the benefit is the intention of the giver" (Seneca, *De Beneficiis* I.V.2). How might benevolent attitudes benefit one? One possibility is that benevolent attitudes (or perhaps one's *knowingly being their object*) can themselves promote one's welfare interests (i.e. things necessary for the pursuit of any kind of life).[10] Another, which I suspect is closer to what Seneca has in mind, is that one's being the object of another's benevolent attitude can itself promote (non-

---

[9] A.D.M. Walker (1980: 43). See also Roberts (2015: 889), for an example in which a "mother is grateful for the gift *to her son*. She does not construe herself as the beneficiary, but him. This kind of case shows that the beneficiary in the triad need not be the subject of the emotion."

[10] For example, "continued life, bodily vigor and health, mental acuity, command of material resources in the form of private property entitlements, adequate nutrition and shelter, security from bodily attack and physical damage to one's person and property, security from theft, fraud, and extortion." (Arneson 2005).

welfare) interests, as might be specified by objective list theories of well-being that include attitude-independent goods.

Notice, however, that even for the "pure" quality of will theorist, it is the *manifestation* of good will that praise is responsive to. Reactive attitudes are standardly understood as reflecting a demand for "*the manifestation of* a certain degree of goodwill or regard on the part of other human beings towards ourselves"; reactive attitudes are "reactions to the quality of others' wills towards us, as *manifested in their behavior*" (Strawson 1962: 84; 83, italics added); "when speaking of *quality of will*…[a]ll I mean, and in my estimation all Strawson meant (e.g. 75, 76, and 83)," writes McKenna (2012: 59) "is that the regard or concern one has toward others (or oneself), and toward the relevant of moral considerations, *as manifested in one's conduct*" (see also e.g. Shoemaker 2013; 2015; Rosen 2015: 77). A slightly different way of putting what is essentially the same point is that reactive attitudes are reactions to the "moral meaning of others' morally significant *actions*" (Macnamara 2011:89). The Senecan idea that gratitude and praise target only benevolent *attitudes*, then, does not appear open to the quality of will theorist. This is because Strawsonians, and other quality of will theorists (Arpaly 2002; 2003; 2006; Arpaly & Schroeder 2014; Markovits 2010), hold that the proper objects of moral responsibility are *manifestations* of quality of will, manifestations *in conduct or behavior*.[11]Because praise and blame are according to her responses to *manifestations* of will, circumstantial moral luck is woven into the fabric of the quality of will theorists' view. The success of nearly any imaginable manifestation of will, especially if by this we mean an overt physical action, depends on the cooperation of my circumstances, which plausibly

---

[11] Perhaps also in emotional *response*, but even this is more than what Seneca allows.

include not only environmental and historical factors, and other persons, but also factors also internal to the agent but external to the psychological features relevant to her 'will'. I return to this issue below. For now the point is only that, while it might occasionally be felicitous to refer to a benevolent attitude (or a benevolent *attempt*) as the benefit, the quality of will theorist's focus on *manifestation* of will warrants distinguishing between benevolent attitudes and the benefits that such attitudes might give rise to.

Next, one can manifest good will without having *acted* beneficently. This follows from the existence of benevolent omissions. For instance, Pat might omit listening (even quietly) to her favorite late-evening radio-show because she knows her sister is a light sleeper and has an important interview in the morning. Strictly speaking, then (*unless* omissions are a species of action) gratitude need not be a response to beneficence, i.e., good action.[12] The benevolence condition will be further precisified, but it will be helpful to do so with the other conditions on the table.

1.2. *The Appreciation Condition*

A mark of S's being grateful to T for ϕ-ing is satisfaction of

---

[12]Smilansky (1997) is right to point out the tendency to overemphasize *positive* displays of benevolence, but the import of his point is nonetheless unduly exaggerated by the presentation of his argument as one for the view that we often have reasons to be grateful for *non-maleficence*. The examples Smilansky introduces are not ones of simple non-maleficence, but non-maleficence *requiring effort and consideration*. For this reason, he is really talking about *benevolent* non-maleficence or benevolent omission. For example, he writes of "the constant vigilance that, say, a driver requires if she is not to harm pedestrians or other vehicles"; "one often keeps the radio down *out of consideration* for ones neighbours" (Smilansky 1997: 589— italics mine).

The Appreciation Condition, i.e. that S be disposed to appreciate or take joy in T's φ-ing.

In many cases, the satisfaction of this condition will consist in the beneficiary's feeling joy, but I leave open the possibility that one can appreciate a benefit without (any disposition to feel) positively valenced affect, like joy. It is not difficult to see, though, why gratitude is considered a pleasurable emotion (and why Bentham for instance includes within his taxonomy of pleasures those of benevolence, or why Martin claims that "gratitude is the normative analogue of the joy we feel when our non-normative hopes are realized"[13]).[14] On the assumption (ubiquitous in the psychological research) that all emotions have either a positive or negative affective valence, it will be trivially true that gratitude is a pleasant emotion. The appreciation condition, however, might conceivably be satisfied if the beneficiary is disposed to appreciate the benefit in the sense of *noticing* and *approving of* the benefit, even if his affective profile is such that he (owing to severe depression, for example) lacks the disposition to have above neutral affective experience (one worthy of the label 'joy') in response to the benefit.

The appreciation condition is satisfied by one's possession of the *disposition* to appreciate or take joy in the benefactor's benevolent action (or omission). One can count as being grateful even when one's occurrent states are not appreciative or joyful. While being pulled over for speeding, Alex will not be appreciating or joyful about much, but he

---

[13] Martin (2013: 125)

[14] Bentham (1781/1988: 36). Nevertheless, as Roberts (2015: 888) notes, to consider "gratitude an essentially pleasant emotion…is fully compatible with someone's finding the *idea* of gratitude unpleasant."

can nevertheless be truly described as grateful to Ben for saving his life, assuming he is disposed to appreciate the action when reflecting on Ben's benevolent deed.

To illustrate that appreciation is indeed a constituent of gratitude, let's consider the ardent egalitarian. Even if her benefactor acts from good will for her, if the egalitarian is given preferential treatment, her lack of appreciation about being so treated might preclude her being grateful.[15] Or supposing that Rex drives Anna's son to baseball practice out of good will, if the drive culminates in a car accident that leaves Anna's son seriously injured, it will be difficult her to be grateful to Rex. In some such cases, the joy of gratitude might be, as it were, crowded out by negative affect, but it's not obvious that this must be the case. This is of course compatible with Anna's having a *reason* to be grateful. The point here is only that Anna will not count as grateful if she is not appreciative of *any* aspect of Rex's driving her son. But, if she isn't grateful at all, I suggest this is because the appreciation condition is a mark of gratitude and this condition is not here satisfied.

It bears noting that we are generally more grateful for benefits we take to be of greater importance to our lives. For instance, I may be grateful to the bus driver who—noticing that my ankle is sprained— drops me off at my doorstep, rather than the closest stop on his route. Yet, I will surely be more grateful to the neighbor who, upon hearing my carbon monoxide alarm, rescues me from a poisonous afternoon nap. This is so because walking a block with crutches does not matter to me *as much* as my life does. That is, in the latter case I receive a much greater benefit, though it is essential that I *take it to be* a much greater benefit. It is conceivable that for some other agent, the normative weight of the benefits will be reversed.

---

[15] Walker (1981: 49)

Additionally, we are much better are noticing that we've received a benefit than that we are at noticing that we've been *benevolently* benefitted. Lacking the ability to apprehend others' attitudes and intentions directly (i.e. unmediated by behavior), the acknowledgment of benefit typically precedes and enables acknowledgment of benevolence. That is, received benefit usually plays an evidentiary role in my being grateful to my benefactor. This is not to say that we draw careful inferences from the one to the other—acts of benevolence are too ubiquitous for this sort of detective work. Still, on the basis of being rescued from the 'carbon-monoxide nap' by my neighbor [receiving a benefit], I will likely be grateful [attributing benevolent motivations to my benefactor], even though I am not certain (and perhaps cannot even be said to *know*) that my neighbor acted from good will. We are not generally so cynical as to require (anything close to) certainty in such cases. Nevertheless, the benefit is a *defeasible* indicator of benevolence. If hours after being rescued by my neighbor, I learned that he woke me *only because* he plans to violently torture me next week (he's been planning it for weeks), I will probably cease to be grateful to him.[16]

I say that the benefit "usually" plays an evidentiary role, because in some cases we may know that one intends to benefit us, without yet knowing what the benefit is. Additionally, I may be grateful to someone for *trying* to benefit me, in which case a *non-benefit* may play an evidentiary role in my being grateful. For instance, if my nephew and his friends genuinely try to make me breakfast in bed when I'm sick, I might be grateful even if the product of their culinary efforts is inedible. In the best cases of gratitude, the

---

[16] I will nonetheless continue to be happy *that someone* woke me (or just, *that* I awoke) from what likely would have been my last nap.

beneficiary will appreciate both S's motivations and the resultant state of affairs (i.e., the benefit).

Our final mark of S's being grateful to T for φ-ing,

The Expression Condition i.e. that on the basis of T's φ-ing, S is disposed to overtly reciprocate.

The grateful person is disposed to express her gratitude. The disposition is to *overtly* reciprocate because typically the beneficiary not only wishes her benefactor well, but wants the benefactor *to know* what the benefit meant for her. In many cases, letting the benefactor so know—by saying *thanks*—is all that is involved in the beneficiary's reciprocating good will. The element of making the gratitude overt, or conspicuous, is noted by Adam Smith (II.III.10), who writes

> What gratitude chiefly desires, is not only to make the benefactor feel pleasure in his turn, but to make him conscious that he meets with this reward on account of his past conduct, to make him pleased with that conduct, and to satisfy him that the person upon whom he bestowed his good offices was not unworthy of them.

While Smith refers to the beneficiary's reciprocation as a "reward", expressions of gratitude, unlike the giving of a reward, need not proceed from an aim to benefit. Smith thinks of expressions of gratitude on the model of intentional action likely because he has in mind gratitude for grand benefits, those that are thought to generate 'debts of gratitude'. Walker too thinks of expressions of gratitude on the model of intentional action in writing that the "grateful person must see his being favored as a reason for wanting to favor another," (1989: 50). But, we can be grateful without having, and certainly without feeling that we have, anything like debt to repay. In some of these cases, we express our gratitude immediately and spontaneously in reply to expressions of good will. I am grateful to the

person at the grocery store who invites me to check-out ahead of him (his shopping cart overflowing with produce, and my basket containing only a case of coconut water). My gratitude is expressed with a smile and 'thanks' in response to his offer, the smile and the utterance both being expressions of gratitude. Neither is done with the intention to benefit, and only one is an action.[17]

While we might genuinely speak of "debts" of gratitude, there is a difference between the reciprocation involved in gratitude and the repayment of a standard debt. Is this an important difference? Not according to those who view gratitude 'externally', e.g. from the standpoint of utility, or with the anthropologist's eye. Those who view it externally hold that the "value of gratitude lies in its promotion of utility, and this end it secures in a wholly external way by supplying *incentives* to beneficence and helpfulness." (Walker 1980: 41). One need not deny that there is utility to belonging to a community that includes expressions of gratitude to reject that the external view adequately captures either the nature or the value of gratitude. Unlike the repayment of a debt (or even the fulfilment of a promise to benefit), gratitude essentially involves sincerity. I can successfully repay a debt regardless of the attitudes I have towards my creditor, but gratitude is partly constituted by the beneficiary's *wanting* to make a return of kindness (at least in part) for its own sake. In light of this latter factor, the character of the reciprocator is apt to partly determine the form or content of the return. In expressing gratitude, one wants to convey

---

[17] Additionally, even when one's expression of gratitude *does* take the form of intentional action (perhaps cognized as the making good on a debt of gratitude), the beneficiary need not express gratitude with the aim of benefitting the benefactor. One reason this is so is that some debts of gratitude find their natural expression in the *projects* we take up in the spirit of the benefit received from the benefactor. See Callard (*forthcoming*) for discussion of a case like this, and for a way of classifying debts of gratitude that sheds light on the different ways in which the *expression condition* is satisfied.

what the benefit *means* to one. The satisfaction of the expression condition for this reason allows for more personality and creativity than that afforded by the repayment of a standard debt. My gas company is owed an exact sum for their services, and no other return, however well-meaning and sincere, will suffice.

Additionally, while the creditor can *demand* that the debt be repaid, when a benefactor appeals to his previous act in order to exact payment, or even to 'guilt' the beneficiary into 'returning the favor', we sometimes think that the *basis* for reciprocating is either undermined or diminished (even if the original benefit *was* performed from good will, though we will naturally doubt this if repayment is demanded) (Berger 1975: 300). Kafka implies a similar view when he writes, "[p]arents who *expect* gratitude from their children (there are some who even insist on it) are like usurers who gladly risk their capital if only they receive interest" (1949: 97, italics added). (Of course, to deny that parents ought to become parents for the sake of repayment is not to deny that filial ingratitude will often be vicious.) The main point here is just that we recognize a nonnegligible difference between the repayment of a debt (/fulfilment of a contract) and the reciprocation of gratitude. The latter has its source in the beneficiary's good will, even if it has a normative (and psychological) weight for the beneficiary such that it is fitting to speak of a 'debt' of gratitude.

The expressive condition helps us distinguish between gratitude and joy that takes as its object another's act of good will. The latter, as Nussbaum, notes need not include (or give rise to) any desire: "[j]oy may inspire no desire, or simply the desire to act in some way expressive of joy—to write, to make love, to listen to the Dvorak Violin Concerto."[18]

---

[18] Nussbaum (2001: 135)

By contrast, the person who is merely joyful at having been benevolently benefitted is in many cases the paradigm of *ingratitude*. It is true that we sometimes call the person that is unhappy with her lot an "ingrate," but when we say this, we are saying not only that they ought to *appreciate* what they have, but that they also have a reason to express their gratitude to their benefactor. For, suppose that the ingrate in question becomes appreciative of what they have—they become happy *that* they have had a wonderful upbringing—yet they display indifference to their parents, teachers, and friends (perhaps out of an undue sense of self-determination). This person may very well be cheerful about their circumstances, but if (as I assume we would) we continue to think of him or her as an exemplar of ingratitude, this is because the expressive condition is not yet met.[19]

Having placed in view the three marks of gratitude— the i) benevolence, ii) appreciation, and iii) expressive conditions— I now argue that our experiences of gratitude reveal our recognition of somewhat paradoxical bounds on the value of benevolence.

2. Benevolence Bound

My overall aim for section 3 is to pose a challenge for quality of will theorists by attending to the dynamics of gratitude. My primary goal is to illustrate the way in which gratitude

---

[19] Hume (Appendix to the Enquiry), for example, takes as definitive of ingratitude the beneficiary's failure to properly reciprocate: "Examine the crime of ingratitude, for instance; which has place, wherever we observe good-will, expressed and known, together with good-offices performed, on the one side, and a return of ill-will or indifference, with ill-offices or neglect on the other." It is conceivable that the magnanimous person will not be grateful for small acts of benevolence, *even if* she is slightly pleased by them and recognizes them to be done from good will. *Ex hypothesi*, her not being grateful will be explained in part by her lacking a desire to make a return of kindness to her benefactor. Some benevolent benefits might be taken to be too minuscule to warrant gratitude.

and praise are susceptible to resultant moral luck, and to argue that intuitive appropriateness of this phenomenon creates trouble for those who wish to reject moral luck outright. By way of stage-setting I first outline two less contentious constraints on the value of benevolence for praise and gratitude. All the constraints in question presuppose that benevolence and praise are scalar in nature, i.e. that they come in degrees. So, a brief word on that is in order first. I draw from the literature on moral worth, since "moral concern" there one way of (though not fully) specifying 'benevolence'.

2.1 The Scalar Nature of Praise

In discussions of moral worth, a contrast is often drawn between someone who does the right thing for the right reasons (/right motive), and someone who performs the same action for non-moral, or amoral, reasons. Kant's self-interested grocer, for instance, may charge customers the same prices as does the dutiful grocer—they both do the right thing— but only the latter's action has moral worth. This point has application beyond Kantian theories of moral motivation, for the core idea is, as Arpaly (2003: 69) puts it, that "[t]wo actions that are equal in moral desirability may be of different moral worth…" and that this difference is explained by a difference in quality of will, however the latter is construed.

The difference between two agents need not be as stark as that between the dutiful and self-interested grocer. Two agents might both act for the right reasons, but with different *degrees* of moral concern or benevolence. To illustrate, imagine first a "diehard" benevolent person, someone who,

> cares so much for her fellow human beings, or for what she takes rightly to be her moral duty to them, that she would act benevolently even if severe depression came upon her and made it hard for her to pay attention to others. [**the diehard benevolent agent**]

<div align="right">(Arpaly 2003: 87)</div>

Next,

> imagine benevolence's fair-weather friend, who acts benevolently as long as no serious problems cloud her mind but whose benevolent deeds would cease, the way some people drop their exercise programs, if there were a serious crisis in her marriage or her job. [**the fair-weather benevolent agent**]   (Arpaly 2003: 87)

Both agents act for the right reasons, but the fair-weather agent's degree of benevolence is far weaker than the diehard agent's, and so, is more easily overridden by morally irrelevant considerations. It is natural to think that intrapersonal moral development consists partly in one's motivationally, affectively, and cognitively fortifying commitment to values one already recognizes. The idea that benevolence or moral concern come in degrees fits this developmental picture. The fair-weather benevolent agent might hope to be more like her diehard peer, and with time and effort she might become so.

Suppose the fair-weather and diehard agents perform the same morally desirable, or right, action. Whose action has more moral worth? That is, who deserves more praise? According to Arpaly, the diehard agent does. Though he might do the right thing for the right reason (i.e. from praiseworthy motives), and so genuinely be worth of some praise, the fair-weather agent nonetheless "deserves less praise than those whose concern for morality or for the well-being of others is more *serious* or *deep*." (Arpaly 2003: 88) This conclusion isn't obvious. One might think that the fair-weather agent will have to *try harder* than the die-hard agent, and for that reason deserves more praise. Not only does she do what is right for the right reasons, she overcomes the strong temptation to do otherwise!

To see why Arpaly takes the diehard agent to be more praiseworthy, consider her statement on what it is to perform an action with moral worth, i.e., to be praiseworthy:

For an agent to be morally praiseworthy for doing the right thing is for her to have done the right thing for the relevant moral reasons –that is, for the reasons for which the action is right (the *right reasons* clause); and an agent is more praiseworthy, other things being equal, the deeper the moral concern that has led to her action (the *concern clause*).

(Arpaly 2003: 85)

Notice that Arpaly's concern-clause states that an agent is more praiseworthy the deeper her moral concern, *other things being equal*. I will soon argue that differences in 'effort expended' can render things relevantly unequal, but for now we should stipulate that no such difference exists between the fair-weather and diehard agent. We can suppose that the fair-weather agent overcame no significant barrier in acting well; there was simply no salient non-moral in competition with the moral reason on which he acted. The point here is that there *may* be cases where two persons both do the right thing for the right reasons, with the same degree of effort, even though one displays better will than the other, and so, is more praiseworthy.

Now for the amendments. First, we should reject the concern clause. That is, we should reject the idea that "an agent is more praiseworthy, other things being equal, the deeper the moral concern that has led to her action (the *concern clause*)" (Arpaly 2003: 84). Depth of concern (or benevolence) plateaus relative to the normative weight on the act-type in question. That is, past a certain threshold, an increase in benevolence fails to contributing to praiseworthiness. This claim is reflected in the intuition that past a certain threshold, an increase in *perceived* benevolence stops contributing to the beneficiary's degree *praise*. Consider the following case of small-scale benevolence:

> While Liv was home visiting her parents, a library book of hers was recalled. As she is out of town and cannot return the book herself, in order to avoid being fined late-fees she asks her friend, Lex, to return the book for her. Lex returns the book in order to help her friend (for the friend's own sake).   [**book return**]

Lex satisfies the right reasons clause and the moral concern clause. She does the right thing for the right reasons. Now, is it true that Lex will be *more* praiseworthy in every world in which she helps Liv from *greater concern* than she does in the last? Upon reflection, the answer is a resounding 'no'. Intuitively, there is a sufficient level of concern or good will that would render Lex *as praiseworthy as she can be* for an action of this caliber. That is, while a much better moral agent, Lex*, would return Liv's book from an almost bottomless well of moral concern—her motivation to return the book would be undermined by far fewer non-moral factors— this seems not to contribute to Lex's* praiseworthiness, nor to detract from Lex's.

In more recent work co-authored with Timothy Schroeder, Arpaly amends her view to accommodate this point, in response to a similar objection from Julia Markovits. There Arpaly and Schroeder (2014: 188) draw a distinction between an agent's quality of will (understood dispositionally), and the quality of will *manifested* in a given action. Adapting a passage from Arpaly and Shroeder (2014: 189) to suit the present example, "[a]n opportunity to assist [a friend by returning a library book] is not typically an occasion for a full display of a powerful commitment to morality. Hence, the strength of desire for the right or the good that is actually manifested by [Lex and Lex*] is the same." There are circumstances in which Lex* but not Lex would, owing to their difference in good will, rise to the occasion. But, these are not the *present* circumstances, and the counterfactual

fact that Lex* would *not* have returned the books in various non-actual circumstances has no apparent bearing on Lex's *actually* deserved praise (cf. Markovits 2010: 213).

To accept this, of course, is to accept that *circumstantially lucky* factors may affect one's worthiness of praise. That is, it is to accept circumstantial moral luck. Is this something the pure quality of will theorist can legitimately accept? Arguably, and in one non-trivial sense, yes. That is, the existence of circumstantial moral luck does not obviously conflict with the following principle that the pure quality of will theorist relies on to determine an agent's praiseworthiness:

Quality of Will Thesis [FOR PRAISE] (QWTp**):**

An agent is worthy of praise only for, and proportion to, goodness of will (/benevolence) as manifested in her action.[20]

This principle, QWTp, takes as given that there *is* a manifestation of good will to be assessed. Contrast a positive thesis of this sort with Nagel's (1981: 58) *Control Condition*.

---

[20]One might want to qualify QWTp such that one can only be praiseworthy for *right (or good)* action. (When moral worth is understood in terms of doing the *right thing* for the right reasons (or on the basis of its right-making features), as it customarily is, this qualification is presupposed (Arpaly 2002; 2003; 2006; Markovits 2010; Sliwa 2016). Such an amendment, though, would arguably require abandoning a pure quality of will view, for to say that one's being praiseworthy depends upon one's doing what is right/good is, at least given standard ways of understanding the rightness or goodness of an action as independent of the agent's epistemic and motivational states, to admit that praiseworthiness is determined by factors other than the manifestation of good will. The view that one can be praiseworthy for doing what is wrong has found support in Haji (2007); Zimmerman (2004); Graham (2010). For simplicity's sake, I only discuss cases of praise for actions that are good/right, leaving aside the question whether or how their goodness/rightness contributes to praiseworthiness.

Control Condition (CC): agents cannot be morally assessed for factors beyond their control.

CC seems to deny that an agent can be praiseworthy for manifesting good will if the circumstances in which the good will is manifested are beyond the agent's control, as they are necessarily are (given the world we live in and kinds of agents we are). By contrast, the idea that luck in circumstances can affect one's praiseworthiness flows from QWTp. It states that *that for which* one is responsible is a certain sort of manifestation, the existence of which of course relies on the circumstances being one way rather than another, and does not require that the circumstances of the manifestation be within one's control (or, for obvious reasons, themselves manifestations of one's will). Since quality of will theorists hold that praise is responsive not simply others' benevolent dispositions, but to the *manifestations* of those dispositions,[21] circumstantially lucky praise is not by itself a problem for them.

Perhaps it can be shown that QWTp is reducible to CC. But if not, the quality of will view is not obviously threatened by the acceptance of circumstantial luck.[22] I will not pursue the question whether CC reduces to, or has its normative basis in, CC (or a further condition

---

[21] What (pre-)dispositions an agent has, to begin with, is also subject to considerable luck. This sort of luck, sometimes called "constitutive luck" (Nagel 1979: 28) is typically accepted by quality of will theorists as simply a feature of responsible human agency, rather than an obstacle to it. See e.g. Arpaly (2003: 169-171; 2006: 31-37). On the compatibilist's default commitment to the possibility of circumstantial and constitutive luck, see Hartman (2016: 2558). Whether one form of moral luck can be accepted without accepting the others is a difficult issue that I here put aside.

[22] The approach I offer on behalf of the quality of will theorist draws on the approach that Russell (2004: 49) advances in arguing that Adam Smith's treatment of moral luck has important advantages over Nagel's.

that both it and CC are reducible to). I turn instead to other, more immediate, grounds for rejecting QWTp.

Instead of comparing agents with different qualities of will, let's compare agents with identical motivational profiles (or a single agent, considered counterfactually), in two slightly different circumstances. Doing so, I propose, strongly suggests that effort itself contributes to moral worth, in opposition to QWTp. Consider the following example:

> After hours of searching for his beloved cat, Allen finds her stuck atop a tree. Caring deeply about his pet, which rarely leaves the living-room, Allen is distressed when he realizes that (owing partly to a back condition) he's in no position to rescue the cat; Allen lives in a sparsely populated country-side, and the telephone has not yet been invented. Fortunately for Allen (and his cat), Rita is at this moment practicing for a marathon, and her run has brought her from the neighboring town to a stone's throw from Allen's yard. Seeing Allen in distress calling out to his cat, and learning about his predicament, Rita rescues the cat from good will. **[cat rescue]**

Now, compare two ways the rescue may have transpired:

> **i. The effortless rescue:** Rita calls the cat's name and the glint of Rita's necklace attracts the cat; the cat climbs down the tree, jumping into Rita's arms.

> **ii. The effortful rescue:** Rita* calls the cat's name and the cat hisses. With the help of a ladder, which Allen is no position to use (back problems), Rita* climbs the tree. After 20 minutes of struggle with the panic-stricken and regularly hissing cat, Rita* makes negligible progress. The cat tightly clutches a medium-sized tree branch. Since the cat shows no sign of loosening its grip on the branch, Allen proposes that Rita* saw off the branch, which Rita* does, bringing the distressed cat to safety.

Assuming that the marks of gratitude are satisfied in "cat rescue", is it not plausible that the effortful rescuer, Rita*, will garner more praise than does the effortless rescuer, Rita? And, putting ourselves in Allen's shoes, would we not feel that Rita* *deserves* more gratitude and praise than does Rita? I think we would. One way to illustrate this is by appeal

to the expression condition. Suppose Allen were to thank the cat rescuer by inviting her to the dinner that he has waiting. Even if this expression of gratitude would not be *unfitting* in response to *Rita's* rescue, it would be significantly more fitting in response to Rita's* rescue.[23] While it might be an excessive way of thanking Rita, it seems like a fitting way to show one's thanks for Rita's* arduous task.

Does "cat rescue" pose problems for the quality of will theorist? One might think it does not. The quality of will theorist already granted that praiseworthiness is affected by circumstantial luck. But given that QWTp states that praise is sensitive to the goodness of will *manifested*, which itself is subject to luck in circumstances, one might think that "cat rescue" is more of the same. Note first that "cat rescue" is unlike paradigmatic instances of circumstantial moral luck, wherein an agent's contingent circumstances present her with a "moral test" (Nagel 1979) that as a matter of the circumstances, which are beyond her control, renders her but not her neighbor (or her counterfactual self) morally culpable for some grave injustice. Small-scale cases of circumstantial luck (like cat rescue) seem to function differently. While Rita* faces a challenge Rita does not, this is unlike the circumstantially lucky challenge of a judge's being offered a bribe. That is, Rita's* challenge does not come in the form of a "moral" test. Yes, Rita* spends more *time* and *energy* helping Allen than does Rita, but time and energy are not what the quality of will

---

[23] On the idea of degrees of fit, see Zimmerman (2015: 55): "we often can and should distinguish between *degrees* of fit. Suppose that Rachel deserves a modest reprimand; a mild one would be too lenient, while a strong one would be too severe. Still, it might be said, neither a mild nor a strong reprimand would be positively *un*fitting; each would be fitting to *some* degree, just not *as* fitting as a modest reprimand would be (whereas it *would* be positively unfitting to praise Rachel)."

theorist takes praise to be responsive to. Recall that QWTp states that an agent is worthy of praise only for, and proportion to, goodness of will (/benevolence) as manifested in her action. To deal with "cat rescue", then, the quality of will theorist seems forced to hold that Rita* manifests more benevolence than does Rita. But our only evidence of this seems to be that Rita* puts in more work or effort. It is not obvious that we have to assume that Rita* somehow 'draws on more of her moral concern', whatever this might mean, in saving the cat. If Rita's* circumstances were such that she was faced with a moral issue Rita does not face, then her circumstances may have required a greater manifestation of benevolence. But on the face of it, what Rita* manifests is more effort. She *does* seem more praiseworthy, but not because her action displays greater moral concern. Rather, she seems more praiseworthy because of the sacrifice characteristic of her rescue.

I submit, a morally concerned (i.e. benevolent) agent is more praiseworthy, other things being equal, the greater the effort expended in his or her expressions of moral concern. Call this the "effort clause". Naturally, fits of effort are not *themselves* praiseworthy, and so, it must be kept in mind that the 'effort clause' is operative alongside other conditions necessary for praise. The effort clause plausibly serves as a contingent *amplifier*, rather than necessary condition, of praise and gratitude.[24] It is plausible that just as the value of benevolence plateaus relative to act-type, so too will the value of effort. Additionally, there will intuitively be various defeaters on the effort-clause. For example, effort will not contribute to deserved praise if it is undertaken for reasons irrelevant to the reason that makes the action right. It might be more *impressive*, overall, if Lex returns Liv's books *by balancing them atop her head all the way to the library*, but she won't merit more

---

[24] On this idea, Gulliford and Morgan (2014: 21); Wood et al. (2008).

gratitude or praise for returning the books in that difficult way if she opts for this mode of return out of, say, curiosity in her balancing abilities.

The effort clause identifies a feature other than benevolence to which praise seems properly responsive, namely effort exerted for the sake of doing what is right. In addition to the effort a benefactor exerts in manifesting good quality of will, praise and gratitude can be sensitive to the *consequences* of manifestations of good will. I argue for this claim, which further limits the normative import of benevolence for praise, in the following section.

2.2 Gratitude Feedback

Taking another to have acted from good will, I have argued, is a mark of gratitude. In order for me to count as grateful to you for some action it is necessary that I perceive you to have acted benevolently. Praiseworthy actions that manifest *more* benevolence ordinarily garner more gratitude and praise than those manifesting less, and the exertion of effort can itself plausibly also contribute to the degree of gratitude and praise a benefactor merits. So far, I have focused on praise at a single time (though occasionally, counterfactually, i.e. in different possible worlds). But something peculiar happens in certain cases of diachronic gratitude. Sometimes gratitude changes across time in a way that seems principled but is not explained by QWTP. I argue that reflection on such cases recommends accepting the

> **Gratitude-feedback thesis:** It is sometimes the case that (i) a subject (S) is grateful at a given time (t_1) to another subject (T) for ϕ-ing, and that (ii) S is at some later time (t_2) *more* grateful to T for having ϕ-ed, even though (iii) no change in T's degree of benevolence has occurred between t_1 and t_2, *and* (iv) S is aware of (iii).

Before discussing the normative significance of this thesis, let's consider whether and in virtue of what, it is descriptively apt. To this end, consider Tim:

> When he is eleven years-old, Tim receives a baroque flute from his aunt and uncle. Tim's aunt and uncle have taken Tim to several baroque performances in the past, which Tim has enjoyed. Although Tim has no antecedent desire to play the flute, he is touched by his aunt and uncle's thoughtfulness, appreciates the gift, and expresses his gratitude with a pair of heartfelt hugs and a few words of thanks. Tim satisfies all the marks of gratitude.

> Years go by without Tim's thinking about the flute—he's been busy with other things, like moving away to boarding school. But on his 19th birthday, after watching Bergman's *The Magic Flute,* Tim decides to play with the instrument that his aunt and uncle gave him years ago. Finding this experience not altogether unpleasant and having been hoping to find a hobby, Tim starts to take flute lessons. To his surprise, he soon displays a real knack for the instrument. After considerable time practicing, and eventually several increasingly important recitals and performances, Tim becomes a skillful flautist, then a virtuoso. By the time he is an adult Tim recognizes flute-playing to be his 'calling'. By the age of 40 Tim is a world-renowned flautist whose craft provides him with a sense of meaning in life he never thought possible as a young man. He is immensely grateful to his aunt and uncle. Thinking of it as a small token of his gratitude, he has just dedicated (and then, mailed) his first recorded album to them.

Call t_1 the time Tim receives the flute. Call t_2 the period from which point Tim takes being a flautist to be his calling and career. At t_1, as we have noted, Tim is grateful to his aunt and uncle for giving him a flute. At t_2 Tim is much *more grateful* to his aunt and uncle. Yet Tim is not, and does not take himself to be, the recipient of any greater degree of benevolence between t_1 and t_2.[25] Given that Tim is aware that his aunt and uncle's expression of benevolence occurs only at t_1, in virtue of what is he more grateful at t_2? It cannot be in virtue of an increase in perceived benevolence. Rather, it is plausibly in virtue of the great joy (or appreciation) that the gift has brought Tim. Tim's appreciation *that* he received the

---

[25] Of course, we have to stipulate that they are not further instrumental in the cultivation of Tim's musical talents (say, by encouraging him to take music lessons).

flute has greatly increased by t_2, and this appreciation, as it were, 'feeds back' into his already existing gratitude *to* his aunt and uncle *for* giving him the flute. This, anyway, is a natural way of understanding how it is that Tim's depth of gratitude increases across time.

The presence of this retroactive feedback is compatible with its being true that after t_1 but long before t_2, Tim comes to better appreciate the benevolence of his aunt and uncle's initial action. That is, Tim is but a child when he receives the gift, and it is not implausible that while he *does* perceive the gift to be benevolently given, insofar as he is not yet in a position look after children, maybe he is not yet in a position to see the manifestation of good will for what it *fully* is. In other words, once he is an adult (or perhaps sooner) Tim may be capable of better seeing the latent value of his aunt and uncle's benevolent motives and intentions, i.e., value that was there all along, but not to be seen by a child's eyes. This phenomenon of "enhanced benevolence-perception" is distinct from gratitude-feedback and poses no problem for the quality of will theorist. To help illustrate the difference between enhanced benevolence-perception and gratitude-feedback we can imagine Tim*, who is like Tim in all respects up to t_1, but who never sees the *Magic Flute*, doesn't take flute lessons, or become a world-renowned flautist. Now a father of two, Tim* may be in a position to better appreciate the value of his aunt and uncle's benevolent gift. Still, it is difficult to imagine him being *as grateful* to his aunt and uncle as is Tim (our world-renowned flautist). While adult Tim's* maturity might enable him to be grateful for features of the gift that he did not appreciate as a child, his adult gratitude *is* responsive to the quality of will. As an adult he might simply be in a better epistemic position to recognize the value present in the manifestation of good will at t_1. His ability to perceive others' benevolence is enhanced in virtue of his maturity. But *what* he is grateful for was

there all along. The same is not true of Tim at t_2. Enhanced benevolence-perception is thus distinct from gratitude-feedback.

My aim is to use the gratitude-feedback thesis to put pressure on the account of praise available to the quality of will theorist. In holding that praise and blame are properly responsive only to manifestations of good and ill will, the quality of will theorist adheres to the Kantian dictum (though not necessarily with a Kantian view of what good will *is*), that a benevolent action is not good "because of what it effects or accomplishes, because of its fitness to attain some proposed end, but only because of its volition." (GM 4:394) As we have seen, the quality of will theorist identifies the proper target of praise not with benevolent dispositions or attitudes, but these attitudes *as manifested* in conduct. The Quality of Will Thesis [FOR PRAISE] (QWTp), recall, states that an agent is worthy of praise only for, and proportion to, goodness of will (/benevolence) as manifested in her action.

The quality of will theorist can therefore accept that Tim's gratitude *at t_1* is appropriate. Gratitude here is responsive to a manifestation of benevolence. But if gratitude is to be responsive *only* to manifestations of benevolence, Tim's gratitude-feedback cannot be warranted. The increased praise at t_2 must be construed as displaying moral sensitivity to morally irrelevant features of the manifestation of benevolence, namely its fruitful consequences.

The quality of will theorist may at this point claim that while the gratitude-feedback might reflect something true of *Tim's psychology*, it has no normative significance. That is, while it might be true that people will tend to be more grateful to their benefactors when their acts of benevolence bear fruit, it does not follow that they should be more grateful.

Perhaps gratitude-feedback is simply an irrational feature of human psychology. If so, there is no normative pressure for Tim to feel more grateful at t_2.

In response, note that there are two distinct claims made in the above rejoinder. The first is that I have not shown

i) that Tim *should* be more grateful at t_2

The second is that I have not shown

ii) that it is *rational* for Tim to more grateful at t_2

Let the sense of "should" in i) be roughly that which entails that Tim would be *criticizable* for *failing* to be more grateful at t_2. Let the sense of "rational" in ii) be that which entails that Tim *has reason* to be grateful at t_2, such that his heightened degree of gratitude at t_2 is fitting. ii) is clearer stronger than i). One can have a reason to feel an emotion, rendering one's having that emotion fitting, without it's being the case that one would be criticizable for not having the emotion. For example, if Jen completes a charity marathon, she might fittingly be proud of herself for having done so, but she will not thereby be criticizable if she fails to feel proud. (To use concepts that are better suited to action: pride is in this case permissible but not required.) Now, the quality of will theorist can accept *neither* i) nor the much weaker, ii). For now, let's stick to ii), the claim that Tim has reason to feel more grateful at t_2, such that his heightened degree of gratitude of at t_2 is fitting (and so, rational).

Insofar as the rise in Tim's praise at t_2 corresponds to nothing pertaining directly to quality of will, the quality of will theorist must hold that Tim gratitude at t_2 is irrational or unfitting. Tim has an attitude, she claims, the strength for which he has no *reason*. A problem for the quality of will theorist is that we can readily specify cases in which

gratitude-feedback does appear to be irrational. But these cases differ in important respects

from the one at hand. Consider, for example,

> Tan, who until t_1 is in all relevant respects identical to Tim. Instead of becoming a virtuoso flautist, however, Tan becomes a veterinarian. On the desk in his office, among various photos and souvenirs, Tan displays the flute he was gifted as a boy. One day after work, noticing that his flute would look better were it painted blue, he takes it to the local painter to have a few coats of paint applied. There another person enters to ask for directions to the museum. Unlike the painter, Tan is able to give her directions, and leaving out details, Tan accompanies her on what becomes the beginning of long and beautiful relationship. Early in this relationship, Tan and his partner developed the habit of playing the lottery, using a combination of both of their 'lucky numbers'. Ten years later Tan and his partner renew their vows and the next day win millions in the lottery. At (t_2) Tan experiences gratitude-feedback; thanks to his aunt and uncle, he met the love of his life. In addition to being part of a fulfilling relationship the value of which he could not as a young man have imagined, Tan and his partner have full financial freedom.

For both Tim and Tan the gift of the flute occupies an important causal role in the coming

to be of their respective happy states at t_2. And, we can suppose that they are equally happy

at t_2. That is, Tim and Tan are equally happy about or appreciative of the aspects of their

adult lives that I have highlighted. Yet, there is an obvious difference between the cases.

The difference may be illustrated by noting that if at t_2 Tan were to expresses towards his

aunt and uncle gratitude of a magnitude similar to that expressed by Tim, we think that

Tan's gratitude was inappropriate. At the very least, we would think that it was far less

appropriate than is *Tim's* at t_2.[26] Tan's gratitude-feedback is inappropriate, I propose,

because there is no 'internal connection' between the benevolence manifested at t_1 and the

state of affairs at t_2. Quality of will theorists, however, lack grounds for distinguishing

---

[26] Though, not inappropriate in the same way as is misplaced *resentment*. Misplaced praise, unlike misplaced blame, does not obviously *wrong* the target of the misplaced attitudes. Still, even if Tan's gratitude isn't *criticizable*—though it's not obvious that it isn't, it might nonetheless be unfitting or irrational.

between Tim and Tan. While they can say that the aunt and uncle deserve praise for their benevolently given gift in both cases, the fruitfulness of a manifestation of benevolence is the wrong sort of thing to alter the appropriateness of praise, and so, Tim's gratitude-feedback is just as irrational as is Tan's.

Although Tim's aunt and uncle did not intend or foresee that Tim would be become a world-renowned flautist, to give a child a gift of a musical instrument is to give him part of the means for cultivating oneself into a musician. Given, also, that this instrument is one that contributes to a type of music to the child has shown interest in, the gift has special roots in the what the beneficiary already cares about. It would have been reasonable for Tim's aunt and uncle to *hope* (and maybe foresee) that Tim would *try* his hand at the flute and perhaps learn, or make some progress toward learning, how to play. The benevolently given gift thus *accommodates* becoming a flautist in a way that it does not accommodate meeting a life partner and winning the lottery. Adapting Williams' distinction between luck that is intrinsic to a *project* vs. luck that is extrinsic to it (Williams 1981: 26), we might say that some luck is intrinsic to a gift, while some is not. Though unintended and unforeseen, the resultant luck present in Tim's life at t_2 is luck that is intrinsic to the gift given at t_1.

Let's look at things from the perspective of the aunt and uncle. Gifts are often settled upon as a result of deliberation, and the gift Tim receives is no exception. The aunt and uncle considered various questions in deciding whether to give him the flute. Would Tim make use of the instrument? Would he feel that they are pressuring him to learn how to play? If they do get him a flute, should it be a wooden or a metal one? Would his parents think the gift was excessive? Would they think they're spoiling him? A good baroque flute isn't inexpensive, and the giving of a not inexpensive gift—perhaps especially one that

carries an implicit suggestion: *learn me*— isn't without its occasions for doubt. Considerations like these pose for the aunt and uncle the question: *are we justified in giving Tim a flute as a gift*? They need not have posed themselves the question in these terms. What matters is that they faced some version of the question, '*should* we give Tim the gift of the flute?'— and that in the events leading to t_1 they answered it in the affirmative.

Let me refer to the intentional action of giving Tim the flute as "the gift". The gift is a manifestation of good will that occurred at t_1. Now, the gift can, and presumably will, be affirmed by a range of people variously related to Tim—fans, sponsors, friends, and Tim himself. To affirm the gift—i.e. to have an attitude of *affirmation* toward it—is to be disposed to have the pleasant thought, 'how good it is that is wasn't otherwise'. Affirmation in this sense is a retrospective attitude. It is the contrary of regret.

In reading Tim's autobiography, Tim's fans might learn of the gift, and registering its importance, they may affirm the gift. Had it not occurred these fans would almost certainly not have Tim's music in their lives, for without the gift Tim would almost certainly not have become a musician—for many of his fans, their favorite musician. In this way the gift has a significance now (at t_2) that it did not have then (at t_1). Like his fans, Tim's aunt and uncle affirm the gift. That is, at t_2, they are happy Tim was given the gift of a flute as a child; they don't wish it were otherwise. But they will naturally affirm it in a manner that importantly differs from that of third-parties, like Tim's fans. The aunt and uncle will affirm it first-personally; they are happy that '*we* did that'. That is, theirs is an instance of what we might call *agent*-affirmation (cf. Williams 1981: 27-9).[27]

---

[27] I put to one side the fact that their thought is first-personally *plural*, along with its possible import for the kind of, possibly *joint*, agency reflected in their agent-affirmation. These are interesting questions, but ones for a separate occasion.

Now, because the significance of the gift is not what it was at t_1, the object of their agent-affirmation is not what it was at t_1. The meaning of the gift is such that after t_2 Tim's aunt and uncle not only gave him a flute that he gratefully received; one can truly say that they introduced him to his vocation. And, it is plausibly in light of the latter fact that Tim has, and takes himself to have, a reason to be grateful to his aunt and uncle that he did not have at t_1 (cf. Wallace 2012: 172). Tim's gratitude after t_2, and his aunt and uncle's agent-affirmation form a pair. Tim's gratitude can be understood as inviting them to *take credit*. That is, we might think that his expressing gratitude is part of what license their agent-affirmation. When an agent is invited to take credit (i.e. to receive praise) in this way, part of what she is presumably invited to do is deepen the way in which she jointly values that for which she is invited to take credit. The invitation may be declined—as it might be if Tim used his musical platform for nefarious ends— but by accepting and thereby adopting a stance of agent-affirmation toward it, one reaffirms the value of the manifestation of good will, now *under the description* that it occupies in the grateful person's life.

As the jacket of his new records indicates, Tim credits his aunt and uncle with introducing him to his vocation. In light of this—though not under the illusion that they are wholly or even largely responsible for his success—their agent-affirmation is colored with pride, both of their nephew and of the gift. Their agent-affirmation is not just, or primarily, a way of 'accepting credit' for the outcome of their gift, but a way of valuing what the gift now means to Tim. This talk of inviting another to take credit is *one* way spelling out the relation between gratitude and agent-affirmation. Though I find it attractive, a proper defense of it would require further elaboration. And as the gratitude-feedback thesis does not depend on this 'invitational model', I leave this defense for another time.

The gratitude-feedback thesis posits the existence of resultant moral luck. Unlike certain cases of circumstantial moral luck, cases of resultant luck are wholly incompatible with QWTp, for the fortunate consequences of a manifestation of will are obviously not preconditions for that manifestation of will, as are the circumstances in which it is based. What's a quality of will theorist to do? Given that the independent appeal of the effort clause, as outlined above, the quality of will theorist is already under some pressure to abandon QWTp. I do not assume, though, that the quality of will theorist at this point *accepts* the effort clause. To assume this would make things too easy.

But, something that might trouble the quality of will theorist is the possibility that resultant moral luck is already presupposed in the idea of a *manifestation* of good will. It takes time, even if only a few seconds, for an ordinary instance of good will to be manifested in action. Now, take any example that the quality of will theorist will accept as a praiseworthy manifestation of good will. If a benevolent *action* (e.g. giving someone the wallet that fell out of their bag) counts as a manifestation of good will, and so, the proper object of praise, then so too, we should count a benevolent *attempt* (e.g. trying to give someone the wallet that fell out of their bag) as a manifestation of good will. One's benevolence is not merely dispositional when one tries and fails to perform a good action. It is manifested. The fact that it does not fulfill the agent's aim seems irrelevant to its *being* a manifestation of good will. But, whether one's benevolent disposition is manifested in a successful action or a mere attempt seems to be a matter of resultant luck. That is, earlier stages of the manifestation of good will—which, according to the quality of will theorist is the proper object of praise (as per QWTp)—ideally have as their consequences the later stages of the manifestation of same benevolent attitudes. But, when we praise a successful

benevolently motivated action, we will be praising what is a *contingent result* of its yet earlier, in progress, manifestation of good will. In other words, our praise will be sensitive to resultant moral luck.

For the above objection to work more would need to be said about the individuation of manifestations of will, which I suspect would not be easy to come to an agreement on. I will not pursue that here. Instead, let's consider why one might be attracted to QWTp, the principle that an agent is worthy of praise only for, and proportion to, goodness of will as manifested in her action. Critics of moral luck typically object that it is "unfair" for something other external to agency to play a role in determining one's worthiness of blame or praise. This is a sensible kind of thing for the quality of will theorist to say too. It should not be a normative primitive that moral responsibility is appropriately responsive only to manifestations of will (perhaps especially given the lack of clarity regarding what a manifestation *is*). On the assumption that QWTp is not reducible to CC, what is the notion of fairness here invoked? The answer to this question should give us an answer to '*why* is Tim's praise at t_2 inappropriate?' that is more informative than 'Tim's praise at t_2 is not proportionate to the quality of will manifested'.

When resultant moral luck is deemed unfair, cases of blame are usually being considered. It is claimed that it is unfair to blame an agent, A for something, X when A could not have prevented X. If X is a product of resultant luck, X is arguably by definition something that A could not have prevented. For, A's relation to X is such that there is another possible agent A* who *is like A in all relevant agential respects*, but does not bring about X. Now, the principle thought to be violated in blaming A for something A could not have prevented is *ought implies can*. This is because it is assumed that one can only be

blamed for doing something wrong, i.e. something that violates a normative demand or requirement. Now, since it was not true of A that A *can* not-X (i.e. cannot have prevented X), assuming ought implies can, it cannot be that X was wrong. But, in that case A cannot be blamed for X because X does not constitute A's violation of a moral requirement.

If the principle that 'ought implies can' is the norm in virtue of which resultantly lucky blame is unfair, a different norm will have to be found to explain why resultant lucky praise is unfair. For, when B is praised for Y it is not the case that Y is something wrong. For this reason, B's being praised for Y does not require that B be able to do not-Y. The principle 'ought implies can' is inapplicable to the justification to praise (Nelkin 2009: 162; cf. Wolf 1980; 1990: 79-81). And indeed, it is not obvious how Tim's praise of his aunt and uncle could be unfair. It is even less obvious that it is unfair *to them*, as it is to those who are unfairly blamed.

The quality of will theorist is in an uncomfortable position. The norm in virtue of which resultantly unlucky blame is thought to be unfair, the 'ought implies can principle' isn't applicable to QWTp. But QWTp is the more specific principle in light of which Tim's praise at t₂ is supposedly inappropriate. In other words, a significant objection thought to face resultant luck leaves intact resultant lucky *praise*. This might not look like a massive blow to the quality of will theorist. But, in the absence of normative grounding for QWTp, she lacks the reason to call Tim's gratitude-feedback inappropriate. Once we grant the existence of resultantly lucky praise, of which Tim's gratitude-feedback is an instance, we will have introduced much more. That is, once we accept that an agent's deserving praise is determined partly by causally downstream features of his expressions of will, our orientation toward this agent and toward a range of more and less (and less and less and

less…) praiseworthy agents will be colored with the thought that the proponent of moral luck was calling the quality of will theorist to acknowledge all along, namely that the agent's quality of will, along with much else that is valued and disvalued in a person's life, is mired with luck. As Paul Russell (2017a: 261) maintains, such a worldview can be disturbing, for "when we consider the various possible ethical trajectories human lives may take—especially those that are ethically unfortunate and problematic—we are confronted with an awareness of our ethical fragility and vulnerability, consistent with conceiving of ourselves as free, responsible ethical agents." Without needing to deny the significance that manifestation of good and ill will has in our lives, this awareness might counsel that we temper our pride with humility and our blame with compassion.

2

Pride, Agential and Otherwise

0. Introduction

Among the things that we might be joyful about, our own actions and traits occupy a special place. After her piano recital, for instance, Alice might feel not simply joy, but also pride. That is, unlike the positive emotion one may experience upon seeing the Northern Lights, Alice's joy evinces a positive assessment, particularly a positive *self*-assessment. The idea that the object of joy must stand in *some* special relation to one's self is present in Hume's observation that "[w]e may feel joy upon being present at a feast, where our senses are regal'd with delicacies of every kind: But 'tis only the master of the feast, beside the same joy, has the additional passion of self-applause." (2.1.IV, p. 290) Hume introduces this example with the aim of precisifying an overly permissive characterization of pride according to which "*everything related to us, which produces pleasure or pain, produces likewise pride or humility.*" (ibid) Baths are pleasant, though seldom the objects of pride. The challenge consists in specifying what the represented "relation to self" must be such that the object of one's joy might also be on object of one's pride. The recital is of course 'related to' the average audience member, qua object of perceptual experience. But the

relation of "experiencing"— even joyfully— is typically *"too slight"* to constitute pride (ibid.)[1] For Alice, however, the recital is not simply something pleasant she experiences; it is *her achievement*. And plausibly she represents it as such. The elevation in self-esteem that Alice experiences in representing this desirable thing as her achievement is plausibly what Alice's pride consists in.

A puzzle arises, however, in the attempt to specify what Alice's pride has in common with various other instances of the emotion pride. Although paradigmatic instances of pride appear to be, as Alice's is, about one's achievements (or at least about products of one's agency), pride can also be felt about *non-actions*, like attractiveness, good memory, or nationality.[2] Unlike achievements and other competently caused actions, one's nationality is not something one can typically be responsible for. Pretheoretically, Alice's type of pride and pride taken in one's traits are the same emotion, but the two seem fundamentally different. One group of philosophers seek to solve this puzzle by denying that this non-agential range of *so-called* pride experiences really *should be* so called.[3] That

---

[1] Of course, the audience member *could* take pride in his attendance, given the right assortment of values and beliefs. Hume is less than ideally clear about a distinction between pride and appropriate pride (though he recognizes this distinction). In saying that "many things, which are too foreign to produce pride, are yet able to give us delight and pleasure" (ibid 291), Hume cannot mean that "foreign things" *cannot* be the objects of pride (unless by 'foreign things' we take him to mean things it would be impossible to coherently represent as bearing positively on oneself, like 'the possibility that "7" is an even number'), but rather that it would be silly to be proud of them (for, he notes that some men *may* feel proud simply about attending the master's feast). See Taylor (1985: 21-23).

[2] This is not even to introduce the complication that it seems that is not only one's *own* achievements that one can feel proud about. Parents feel proud of their children; teachers of their students; sports fans of their home teams, and so on. For this reason, we are doubly blocked from characterizing pride as joy felt about one's own achievements: on the face of it, the object of pride need not be one's own achievement, and it need not be an achievement (or agential) *at all*.

[3] Solomon (1976: 345) writes, for instance, "[t]he key to the emotion pride is that it is about our achievements in the world." In pride, "One takes responsibility (in praise) for his own

is, on the "restrictive" view of pride, the emotion involves viewing an object as expressive of one's agency and something for which one is thereby *morally responsible*. On this view, to be proud of one's natural good looks is either to misrepresent those good looks as products of one's agency, or to mislabel one's *happiness that* one is naturally good-looking. In contrast to this restrictive strategy, others adopt a "disjunctive" account of pride's objects. On the disjunctive view of pride there are really two distinct emotions called "pride": a moral emotion and a non-moral one. [4] Although the disjunctive view accommodates pride for traits and other non-actions, it fails to explain the sense we have that there is unity to our range of pride emotions. Both the narrow and disjunctive views of pride come with costs. For example, one sacrifices breadth for unity, and the other unity for breadth. But, no such sacrifices need be made. A capacious and unified view of pride's objects is available, and its outline and defense sheds valuable light on the relationship between agency and practical identity. So I argue.

The aim of this paper is thus to provide a novel analysis of the nature of pride, by first providing an account of its content, i.e., what the emotion is *about* such that all instances of pride qualify as instances of the same emotion-type. This analysis will provide us with of pride's "attribution conditions," that is, those conditions that must be met to

---

work" and the emotion accordingly includes the desire "'[t]o be recognized for what one has accomplished." (ibid: 346) And, according to Kristjánsson (2002: 104), the emotion pride "differs from mere joy in that the former emotion attributes (some) responsibility for the achievement to the subject." While Kristjánsson does use the label "pride₂" to refer to the attitude that a proponent of "Black pride" might feel, "pride₂" amounts to *self-respect* on Kristjánsson's view. Pride, the emotion, requires that one see oneself as *responsible*, to some degree, for the object of pride (125).

[4] Peter Strawson and Strawsonians generally, adopt this approach. Insofar as pride is paradigmatically a self-reactive attitude of moral praise. In the psychological literature, this view is can be found in Lewis (2000) and Ekman (2003).

intelligibly attribute the emotion to someone. I proceed as follows. In section 1, I set the stage by disentangling various sense of pride. In section 2 I provide an account of "agential pride", outlining why this facet of pride encompasses but is not exhausted by (what is sometimes calls) "achievement-pride" (it also goes by the names 'beta pride' and 'authentic pride', labels which I put aside). In section 3, I provide an account of what I call "non-agential pride." Jointly, these facets of pride provide a picture of practical identity reflecting that, as agents, we are more than what we do.

While there is a rough sense in which "agential" and "non-agential" pride corresponds to what social psychologists sometimes refer to as "achievement-pride" and "hubristic pride", respectively, my division is importantly distinct. There is a tendency to construe 'non-agential' pride as morally suspect; it is sometimes called "hubristic" or "self-aggrandizing" pride (Tracy and Robins 2004; 2007). (In this way, it has the status that envy does according to those who think that emotion is inherently vicious.) This is a flawed way to characterize the phenomenon in question, for it runs together what should be kept apart, namely pride's attribution conditions and its fittingness conditions. On my view, agential pride and non-agential pride—which are roughly the positive analogues of *guilt* and *shame*, respectively—both admit of instances that are fitting, and other instances that are not.

The tendency to treat trait-directed pride as hubristic stems, I suspect, from the ease with which trait-directed pride is confounded with the *trait of* pride, which is often understood as a vice.

1. Proud Persons and Pride Feelings

Some things called "pride" fall outside of this paper's purview. I am interested in pride, the emotion. "Pride" can also refer to a trait of character, one frequently taken to be a vice. Pride qua vice is hotly condemned within Christian doctrine.[5] According to *Proverbs* 29:23, for example, "[a] man's pride will bring him low, But a humble spirit will obtain honor." The "proud person" is one generally disposed towards arrogance, conceit, self-aggrandizement, and dogmatism (for to be open to altering one's views is to admit fallibility).[6] At the extreme, "the "proud person repudiates bondage to human need…the vice of pride [is] the idea that one can and should achieve godlike omnipotence, removing all passivity and need" (Nussbaum 2001: 527). When the proud person's godly status goes ignored, his pride can be "wounded".[7] While this use of the word "proud" is somewhat archaic (at least within colloquial English), we do speak of 'big egos' (which, as the expression goes, sometimes are *bruised*). These more or less correspond to the 'proud person'.

It should not be assumed that one's having the trait of 'pride' requires conceit over *all* (or even *most*) other persons. The trait can be localized. It is partly for this reason that, even when considered as a trait of character, pride is not always vicious; sometimes it can be silly, even endearing (see Taylor 1985: 47). This pair of points is illustrated in the pride

---

[5] See also, *Habakkuk* 2:4: "Behold, as for the proud one, His soul is not right within him; But the righteous will live by his faith"; *Psalm* 40:4: "How blessed is the man who has made the LORD his trust, And has not turned to the proud, nor to those who lapse into falsehood".

[6] On the connection between pride and dogmatism, see Roberts (2009).

[7] Indeed, one advantage Ben Franklin notes about extirpating his (vice of) pride is the experience of "less mortification when [he] was found to be in the wrong." Roberts (2009:121)

of Proust's Françoise. There we find a role-specific and morally non-vicious example of the trait. Marcel describes Françoise as exhibiting "a sort of pride in not allowing anyone else to come near [his] aunt, preferring, when she was herself ill, to get out of bed and to administer the Vichy water in person, rather than to concede to the kitchen-maid the right of entry into her mistress's presence" (Proust 1922: 168). Even if the paradigmatic "proud person"— e.g. Chaucer's Simkin, "as proud as any peacock"(1.3926)— is a disagreeable (because a *generally belittling)* type, one's possession of the trait "pride" is presumably compatible with being, on the whole, a decent person.

I am leaving several threads loose. This is *okay*, however. The "proud person" is not my topic, and neither is the character trait of pride. I need not arrive at much precision on either of these for the following reason. The feeling of neither "agential" nor "non-agential" pride entails that the person feeling it has the trait called pride. One's feeling proud of his son's achievement, for example, is wholly compatible with expressing humility where appropriate (say, when speaking to a specialist). Thus, while the "proud person" might be contrasted with the "humble person", the *emotion* of pride is better contrasted with guilt and shame. [8] I suspect that Tracy and Robins' (2004; 2007) categorizing one facet of pride "hubristic" derives in part from pride's association with the narcissism of 'the proud person'. In reality, however, whether one has a 'big ego' is a matter separate from his occasionally experiencing agential or non-agential pride.

Having put the trait aside, I turn now to agential pride.

---

[8] Specifically, agential pride has as its natural counterpart *guilt*, and non-agential pride, shame.

2. Agential Pride

Frederick Banting felt proud about discovering insulin, we may suppose. This can be intelligibly supposed because one type of pride is pride felt about things *done*, especially one's achievements. Banting's discovery of insulin—a life-saving drug for Type 1 diabetics —presumably qualifies as an achievement, where an "achievement" is a difficult and competently caused attainment of a goal (Bradford 2015). Since this gloss on "achievement" is silent on the goodness of the goal, vicious acts may qualify as achievements. Even if, upon consideration, we elect to reserve the term "achievement" for actions that are *at least* morally permissible, one can nevertheless feel proud about one's wrongful, but 'difficult and competently caused attainment of some goal'. Leopold and Loeb, for example, were at least initially proud of having kidnapped and murdered Bobby Franks, an undertaking they envisioned as both a "perfect crime" and an expression of their imagined exemption from societal norms. "Leopold", according to a notorious study of the case, "seemed proud of his crime, telling a *Tribune* reporter, "Why, we even rehearsed the kidnapping at least three times, carrying it through in all details, lacking only the boy we were to kidnap and kill""(Higdon 1999: 126). Regardless of whether it is fitting, the pride experienced by Leopold is no less agential than Banting's.

In what, then, does agential pride consist? I propose that,

in feeling **AGENTIAL PRIDE** for an action A, a person P views an action of his/hers [or of someone with whom P identifies, e.g. a child or parent or mentee] to be something valuable that reflects positively on P, and in virtue of P's relation to A, P feels joy about A.

First, in saying that the subject of pride "views" valuable her achievement, I take viewing to be cognitively weaker than judging. "Appraising" is preferable to judging, but there are

many appraisal-theories of emotion to choose from, and 'viewing' moreover enjoys the benefit of being the more intuitive (because non-technical) term. I do not provide an analysis of what this attitude is, but I opt for something weaker than judging for two reasons. First, one might feel pride against one's judgment.

> I may find myself getting swept up in the patriotism following the 2001 attacks on the World Trade Center and Pentagon and so come to be proud of the American response—not merely in the bravery of the rescue personnel but also of the quick rout of Taliban forces in Afghanistan. Nonetheless, I may also simultaneously be a pacifist and so judge the use of military force of any sort to be objectionable. In such a case, there is a rational conflict between my judgment and my pride. (Helm 2010: 102)

I do not want to rule out cases in which one's evaluative judgments *do* undergird one's feeling pride, but such judgments appear not to be necessary.

My second reason for saying that the achievement is merely *viewed* (and not judged) valuable is that I do not want to exclude *a priori* the view that the pride of humans and non-human animals belongs to the same emotion type. Cat owners, for instance, will be familiar with the pride-like behaviour a cat can display in delivering to one's doorstep her prey (see Prinz 2004: 28-9). I do not commit myself to the view that cats do (or do not) feel pride. Given that it is a self-conscious or reflexive emotion, pride presumably requires *some* (or at least, *something like*) a concept of selfhood. It is largely an empirical question whether this can be attributed to cats. But, whatever concept of selfhood (or non-conceptual capacity for self-thought) is necessary for pride, it presumably need not be very sophisticated. My niece (who is 12 months old) probably does not have a rich concept of selfhood, but I submit, she feels and displays pride when celebrated for taking (baby) steps.

A comment about "viewing" and self-reference is in order. Pride is paradigmatically a *first-personal* attitude, which is to say that it is an attitude by which one

views an action as reflecting positively on oneself. It is further essential to pride of this sort, however, that the agent view herself positively in a particular way, namely through a *de se* mode of presentation. Self-praise is for this reason non-identical to pride (even if we exclude second-personal pride, which is non-controversially distinct from self-praise).[9] Unlike self-praise, pride is essentially immune to error through misidentification.[10] To attribute pride for an achievement to Alice at some time t_1 is, in other words, to deny the possibility that Alice can sincerely ask at t_1 questions like: "I feel proud about that agent's achievement, but is that agent me?" By contrast, there may be cases of self-praise that are epistemically akin to cases of admiration or gratitude, one's in which one views oneself *de re* as praiseworthy (or responsible for some valuable action), but owing to self-misidentification, one's view of oneself fails to translate to pride. For example, suppose that thirty years from now, actor and former firefighter, Steve Buscemi, by then in his eighties and unfortunately suffering from Alzheimer's, is watching footage of himself at work after the 9/11 attacks. Having forgotten much of his past, he might admire (and so, third-personally praise) 'that firefighter' without realizing that that firefighter is he (—in case one doubts whether there's enough psychological similarity for personal identity to be preserved, we can modify the case such that Buscemi fails to remember *only* that he helped at ground zero; he's otherwise (i.e. presumably, essentially) the same person).

---

[9] Although this point may strike one as too obvious to mention, there are views of guilt in terms of self-blame that imply that guilt is susceptible to error to through misidentification relative to the first-person pronoun. But, just as self-praise is non-identical to pride, so too self-blame and guilt are psychologically distinct.

[10] I am of course drawing these insights from Shoemaker (1968).

Although this is a case in which Buscemi *views* some action as reflecting positively on *himself*, he views himself from the wrong mode of presentation for pride.[11]

Next, the most robust way to view an action as "something valuable that reflects positively" on oneself is to view the action as something for which one as *agentially responsible*. In paradigmatic cases of agential pride, one represents the action as "reflecting on oneself" *by* representing it as something for which one is responsible. Responsible, how? We sometimes bring about desirable effects non-competently, and the non-competence of one's causal role in such cases typically functions as a defeater on pride (presumably because it undermines the *praiseworthiness* of the behaviour). In one episode of *Curb Your Enthusiasm*, Larry David trips on his inanely long shoelaces, and in so doing accidently 'tackles' a passenger who is drunken and abusive (to the flight attendant).[12] The joke is that Larry, being not a wholly honest man, accepts credit for 'being a hero'. The responsibility in question need not be *moral* responsibility, however, even it is true that people tend to view morally relevant attributes as especially integral to personal identity (Strohminger and Nichols 2014). Unlike guilt, which invites moralized analyses of its contents, we readily accept that in feeling proud Alice may view herself responsible for her extraordinary piano performance without supposing that she must view herself *morally* responsible. The concept of "agential responsibility" is therefore narrower than mere causal responsibility but broader than moral responsibility, presumably also including the

---

[11] Or rather, it is the wrong mode of presentation for first-personal pride. This consideration suffices for concluding that self-praise is non-identical to pride. Buscemi might, however, feel *vicarious* pride (about 'that firefighter') in virtue of his being a New Yorker.
[12] At 2:12: https://www.youtube.com/watch?v=Si9gZc0vouE (Season 8, episode 6 "The Hero")

responsibility that an athlete takes for an exemplary performance (Russell 2008; Wolf 2015).

Agential pride cannot be restricted to one's *own* action because agential pride can be vicarious. For this reason, it is inessential to agential pride than one represents oneself as *responsible* for the action that figures in one's pride. Alice's piano teacher (Pia) might feel proud of Alice for Alice's flawless performance of Chopin's Nocturnes. In this case, Pia *is* proud about an action, though not her own. While we may be tempted to say that Pia is proud about *having taught Alice in such a way that enabled Alice* to […], we can put this temptation aside by noting that Alice's father and mother *too* might feel proud of Alice for her performance. Mother and father, however, never taught anyone how to play piano, nor is it plausible that the content of their emotions is better described as pride 'about having registered Alice for piano lessons' [R]. R was true before the piano recital. Both Pia and Alice's parents may well take some credit for Alice's achievements—they have, after all, been nurturing and caring guardians of her flourishing—but note that one can also be proud of the actions of one's parents and (great-) grandparents. In so doing one need not (and barring abnormal beliefs, usually will not) take agential credit for those actions. Still, in feeling second-personal agential pride one represents the agent's action as reflecting positively on the agent of whom one is proud, *and on* oneself, *by identifying with* the agent. The norms governing identification of this sort are plausibly those governing our various relationships; it is part of the ideal of (perhaps, close) friendship that one is proud of one's friend for his achievements. To be proud of another for some action is (at least if the other person is living) to represent oneself as standing in a *supportive* relation to that person. For this reason the abusive parent plausibly lacks the 'standing' to be proud of the now-

successful child, and we will not be surprised if she deflects his pride. Second-personal pride relies on relationship-dependent reasons. The parent-child relationship *ideally* provides the parent with reasons to be proud of the child for her achievements, but because this relation is (like that of friendship) a normative and not biological relation, the existence of such reasons requires that the parent meet the standard of (good) parenthood.

In addition to relationship-dependent reasons for pride, there are membership-dependent reasons for pride. In light of the latter we can make sense of the core case of pride of *being* someway rather than for doing something.

3. Non-Agential Pride

In celebration of the 65th anniversary of India's independence, a widely circulated English-language Indian newspaper asked the youth among its readers why they are proud *to be* Indian. [13] Some of the responses pointed to very general reasons for pride, while simultaneously maintaining that justification is not of prime significance in this area. Govind Pareek, of Candlewick Public School in Jaipur, for instance, writes, "I am proud to be an Indian because of its rich tradition and heritage…I have got tons of reasons to love my country, but the most important thing is I love my country and I need no reason for it." Aspects of the idiosyncratic and (at least, prima facie) non-contrastive character of reasons for pride are captured in one student's attention to Indian's strengths in spacecraft technology. While we might condemn the person who feels guilty about some of his

---

[13] "Proud to be an Indian," *The Hindu*, August 13, 2012: http://www.thehindu.com/features/kids/proud-to-be-an-indian/article3762135.ece

wrongs but not others, there's something of an *elective* character to the objects of pride. I return to this point below. For now, consider an example of what I mean by idiosyncratic national pride (which will also illustrate the way in which reasons for pride may be non-contrastive). Sri Sagana of Coimbatore writes,

> I am proud to be Indian because of our country's achievements in space and missile technology. India has launched a number of satellites in space for helping us in weather forecast, communications, medical research, and education. The missiles such as Agni, Prithvi strengthened our nation's security and ensured a place in the elite group of the world's powerful countries. Moreover, our scientists have lifted India's pride by placing our Tri-colour on the moon and are now looking for landing in the planet Mars.

The idiosyncrasy of Sri's pride is not its uniqueness to *her*; surely other Indians are proud of their nation's achievements in space technology. Still, that this is Sri's reason for national pride—and not, say, India's store of talented young writers, like Mistry, Thayil, Desai, Ghosh, Adiga, etc.— is expressive of particularities of Sri's interests. These idiosyncrasies, however, are normatively kosher. That is, *it's fine* that Sri is not proud to be Indian in virtue of India's great writers. Similarly one's indifference to (though not *derision* of) national space achievements would presumably not constitute a normative failure of any sort. Contrast this with indifference to the war crimes of one's nation.[14]

  Secondly, note also that each of the above reasons is, at least on the face of it, a *non-contrastive* reason for Indian pride. That is, although Sri Sagana is proud to be Indian in virtue of "national achievements in space technology" (AST), AST cannot be—and in fact is not—Sri's reason to be proud of being Indian, *rather than* Russian or American. As

---

[14] The point that pride's objects have an elective character should not be exaggerated, and is compatible with the observation that it might be both unfitting and inappropriate to fail to feel proud in some circumstances.

Sri herself recognizes, there are other nations that have sent probes to the Moon's surface; in fact, relative to that of the USSR and USA, India's contribution to lunar exploration is quite small. I say this not to call into question the appropriateness of Sri's pride (and not only because I leave aside questions of axiology). Sri is Indian and not Russian. Accordingly, Russia's achievements cannot *belong to her* in the way that India's can.[15] Similarly, one need not delude himself into thinking his child is *the best child* in order to be a proud parent (fortunately so for siblings). Perhaps, however, Sri's reason for pride has two parts: AST *and* that "Sri is Indian" ([I]), generating AST[I]. Even if, on its own, [I] is not obviously a reason for pride, AST[I] may nonetheless be Sri's reason for feeling proud of being Indian, rather than feeling proud of Russian. This response provides an avenue for preserving the thesis that reasons generally are contrastive, such that one's reasons for pride are reasons to be proud about x rather than y.[16] There is some oddness, however, in claiming that AST[I] is Sri's reason for feeling proud to be Indian rather than Russian. After all, it

---

[15] Russia's achievements, however, could in an intergalactic (and not international) context conceivably provide Sri reasons for Earth-pride. (The achievements of the "USSR", of course, are not Russia's alone.)

[16] I draw here from Snedegar (2008). Although it leaves out discussion of "reasons to feel", the generality of Snedegar's view (being about reasons, in general) seems to commit to the view that reasons to feel, if there are any, are contrastive. For, "[o]n [his] theory, called *contrastivism* about reasons, reasons are always reasons for one thing rather than another, or more generally reasons for one thing out of a certain set of alternatives, instead of reasons for things *simpliciter*, as has been traditionally assumed." (In passing and on different grounds, Sinnott-Armstrong also commits himself to the view that reasons to feel must be contrastive. See Sinnott-Armstrong (2008). It is possible for Snedegar to nonetheless hold that, for certain emotion types (or tokens), one (appropriately) feels an emotion without a reason to feel. If so, he would have a companion in Govind, who, you will recall, writes that "*the most important thing* is I love my country and I need no reason for it" (italics added).

is impossible for there to be a reason *for Sri* to feel proud of being Russian (at least in the same way that she has a reason for her current Indian-pride).[17]

Fortunately, this issue need not be resolved here. What matters at present is that Sri is "proud to be Indian because of [her] country's achievements in space and missile technology." Now, *being* Indian is not something one does, and it is not something for which one is agentially responsible (in the way that one would be for, say, failing to feed one's pet).[18] Sri neither chose nor agreed to have Indian parents. Yet, one can, as Sri does, feel proud of being Indian. Since one can feel proud about *being* X, and not only about Φ-ing, we should make room for a form of non-agential pride. This claims requires some argument, which will also serve to clarify what the distinction between agential and non-agential pride comes to.

First, isn't my category of "second-personal agential-pride" really just a form of non-agential pride? When Alice's father feels proud of Alice, he is not proud of *doing* anything. The performance, after all, is Alice's and not Dad's. I acknowledge that Dad is not proud about his *own* achievement. But, the content of his pride includes centrally an achievement, namely Alice's flawless performance of Chopin's Nocturnes. Now, achievements include, essentially and necessarily, actions. Actions are exemplary

---

[17] At least, on the assumption that "Sri" designates rigidly and 'being of Indian descent' is essential to Sri. (I deny neither that Sri might someday acquire Russian citizenship and take pride in Russia's achievements qua Russian (citizen), *nor* that Sri might one day, for whatever reason, disavow his Indian-ness, and so, extinguish reason AST[I] by symbolically rejecting [I].)

[18] Of course, if belonging to some nation (N) can give you certain obligations (moral or political), there is a sense in which being a member of N is something for which you are responsible. That is, one's group identity may provide a group-member with responsibilities, e.g., the responsibility to pay reparations for the group's past wrongs.

expressions of agency. For this reason, even though Alice's performance is not an expression of Dad's agency, Dad's pride is agential. And, *because* this agential pride is not about something that is an expression of Dad's agency, it is second-personal agential-pride. (It remains so even if Alice herself fails to take pride in the performance; Dad's pride derives not from Alice's pride, but from her achievement, coupled with his identification with his daughter).

One might worry that this response creates a more serious problem, from the other direction. That is, it now may appear that non-agential pride collapses into agential pride. After all, Sri's pride, recall, *is* in a very important way about India's achievements (specifically in space technology). Why, then, not categorize Sri's pride as kind of second-personal agential-pride? My answer cannot be that India is not a person. If there is an achievement in question—and there is—Sri's pride can plausibly derive from the agency of those responsible for that achievement, namely certain members of the "Indian Space Research Organisation". We can even suppose that Sri knows who they are; she can put a face (or several) on India's achievements. Why then is Sri's pride not simply an example of (second-personal) agential-pride?

Well, *it could be*. This is to say, one might feel proud of the members of the "Indian Space Research Organisation" for planting an Indian-flag on the moon, etc. But, this is unlikely to be an adequate description of Sri's attitude. Many instances of non-agential pride *can* be justified by appeal to achievements, but this does not mean that in feeling proud to be Indian (or French, or American), that the content of one's attitudes include particular achievements that could provide warrant for such pride.

For a clear way to see that this is so, consider instances of non-agential pride that are incapable of being justified by appeal to achievement: one's pride about being fair-skinned, or tall, or pretty, or possessed of a good memory.[19] Foster Wallace (1996: 220) describes a not unrecognizable character when he writes, "[t]he M.P. was…broad and very muscular, and proud of his muscles, going shirtless whenever possible." Following Isenberg, let me call features like tallness and beauty "gifts", in contrast to "achievements" (e.g., discovering insulin, learning Latin, etc.), which by definition, "are the result of purpose and effort". This distinction between gifts and achievements, Isenberg (1949: 368) notes, invites the following pair of thoughts: "I am in no way responsible for the good looks I may have inherited from my mother or the vigor transmitted by my father; so I am said to "deserve no credit" for them." One may expect Isenberg to continue by saying something like the following: 'in feeling proud about something desirable, I thereby *take credit*, or claim responsibility, for that thing. But, I *can't* rightfully take credit for *gifts*—they're simply 'given' —so, when one sincerely says, e.g., 'I'm proud of my looks', he either a) mistakenly takes credit for a gift, in which case his pride is *inappropriate*; or b) he does not, in which, his attitude is *not pride*, but joy perhaps.'

Instead of making the above point, which I think reflects an intuitive line of thought, Isenberg says something of interest. He claims that at a certain level of analysis *all* objects of pride are 'gifts'. For this reason, although his vigor and good looks *were* inherited, there is nonetheless a certain sense in which he can appropriately take pride in them. Isenberg's articulation of this point deserves quotation en bloc:

---

[19] I do not deny that these features *may* be achievements (some of them more conceivably than others), but that being proud about them implies that they are taken to be products of agency.

it must be the intention of this remark [i.e., that he "deserve no credit" for his gifts] not to deprive me of all credit whatsoever but to insist only that such credit should be *less* than the credit accorded for traits of character or acts of will, and should be *different* in quality, as honor, reverence, and admiration are different one from another. For in a more radical sense of the word gift, is not everything in the end a gift? Who gave me talent or kindness or heart, magnanimity, courage, the capacity for self-discipline and self-development? Not I myself. The aptitude for learning is not learned, the ability to acquire is not acquired. My acts are dependent on my will as my stature, for instance, is not. But my will—what is that dependent on? Some people are endowed with the desire for improvement as others with beauty or strength of body; and one is not at bottom more praiseworthy than the other *if it is a question of ultimate responsibility for what one is and does*. It is true that my character has been more deeply affected than my appearance has by decisions which I have made in the past; but does that explain why it should be more seriously praised or blamed?

(Isenberg 1980: 368-9)

Even if natural qualities and achievements are both grounded in features about ourselves for which we are not *ultimately* responsible, we are not mistaken, Isenberg thinks, in according special value to those goods we recognize as expressions of will. Isenberg anticipates one aspect of P.F. Strawson's naturalistic compatibilism in saying that "The will is praiseworthy and blameworthy where the torso is not, not because it had something more to do with creating itself but because the will is the will and the torso is the torso (Isenberg 1980: 369). Causally determined though both the torso and will may be, we nonetheless respond to expressions of the latter in a distinctive way. For Strawson (2003: 83) our attitudes of praise and blame "are *essentially* reactions to the quality of other's will towards us, as manifested in their behaviour: to their good or ill will or indifference or lack of concern". In spite of their similarities, Isenberg's stance differs in a *very* important way from that of quality of will theorists (e.g. Strawsonians). Isenberg thinks gifts (including natural qualities like tallness and beauty) may *also* be something for which we can take credit, and for which (he seems to imply) *appropriately* feel pride. For the Strawsonian, by

contrast, gifts like looks and height fail to reflect the quality of one's will, and for that reason pride felt about them is bound to be inappropriate.[20] And yet, in saying that that the credit accorded beauty is lesser *and* different from that accorded good will, Isenberg recognizes a distinction between two facets of pride.[21]

Given the special responsibility-assigning function that pride *qua* reactive attitude occupies for the Strawsonian, he is forced to say one of two things. Either non-agential pride is a distinct emotion from 'agential (moral) pride', *or* 'non-agential pride' consists in inappropriately representing some 'gift' as expressive of one's quality of will. The first option is unattractive because it requires individuating emotions in terms of their *actual (i.e. non-formal)* objects. We would never deny that one's belief that, say, 'Texas is in Australia' *really is* a belief. That it represents some content *as true* is what matters. The second option is unattractive because it requires attributing to pride a highly determinate formal object (and by extension, correctness conditions), one that many (perhaps most) people who have felt pride will be mistaken about (i.e., everyone who believes that some instance of non-agential pride is fitting). Given that the Strawsonian's naturalistic compatibilism is supposed to be based on a variety of emotions *as we find them* (i.e.,

[20] Watson's use of parentheses suggests that he has this type of point in mind when he writes that the attitudes of interest to Strawson include "(some kinds of) pride." Watson (2004: 220).

[21] Isenberg, nevertheless, incorrectly draws this distinction as one between "natural" and "moral" pride. Incorrectly, because non-agential pride need not be about natural features, and agential pride need not be about moral agency. The idea that we 'take credit' for non-agential qualities is present also in Lazarus' claim that pride consists in the "[e]nhancement of one's ego-identity by taking credit for a valued object or achievement, either one's own or that of some group with whom we identify." Lazarus (1991: 122)

pretheoretically), the formal objects of these emotions cannot be tailor-made to fit a quality of will theory.[22]

Putting aside questions about its possible fittingness conditions, how are we even to distinguish what I call 'non-agential pride'—pride one feels in some non-agential property (e.g., beauty; nationality)— from simply taking pleasure in one's having some property or feature? To see the challenge, consider that '**listening to R.E.M at some time, t_i.**' [LREM] is a property of me, and that I take pleasure in LREM. That is, I am enjoying listening to an R.E.M. song, "The Great Beyond", suppose. Now, it is *not* true that I feel proud of LREM. Given that I need not take the object of my pride to be the product of my agency, what can distinguish LREM from non-agential properties that I *do* take pride in? The answer, in short, is that the non-agential objects of pride are properties that I take to be central to myself. If you like, they are features of my practical identity. Although the supermodel's beauty *is* a gift—and so, like the rainbow that she enjoys, is not a product of her agency— in viewing her beauty as something desirable that is essential to her personhood, she might thereby take pride in her beauty but not in the rainbow, even though she is happy about both. The supermodel's beauty is related to her in a way that LREM is not related to me. It is, *in her view,* an essential (or at least, a *very important)* feature of who she is. (I like R.E.M., but I doubt that my self-identity would be shattered if I could not listen to their music).

More generally,

---

[22] This is not even to introduce the puzzle of restricting "quality of will" in a way that plausibly excludes (the appropriateness of) e.g., the amoralist's pride in some great aesthetic achievement. That is, the formal object of pride must be *very finely* tailor-made by the Strawsonian; I fear it won't fit most of us.

in feeling **NON-AGENTIAL PRIDE** for some property T, a person P views T as something valuable [/desirable] that is expressive of P's true self [/practical identity], and in virtue of so viewing T's relation to P, P takes joy in T's relation to P.

The person who feels non-agential pride represents the desirable property as something that is expressive of her true self. If the reader prefers, we can say that P views T as a feature of her practical identity, or as essential to her personhood. Since one need not have a theory of either essences (or true selves or practical identities) in order to feel proud about non-agential properties, all that matters is that P takes the valued property in question to be integral to who she is, in a way that one does not ordinarily view one's trivial features, e.g., finger-nail length, distance from the Equator, number of eyebrow hairs, etc.[23]

Given that non-agential pride is not about actions, we might wonder whether it too involves states of joy. Since agential pride is about *expressions* of agency (like achievements) it is not mysterious why it should include an *episode* of elevation in one's self-estimation. How, though, are we to understand the corresponding episode of joy—if there is one— when it comes to non-agential pride? The worry underlying this question consists in the following thought: in contrast to one's reasons for agential pride, the *stability* of one's non-agential reasons for pride renders it non-obvious why one should feel non-agentially proud at *this* moment rather than another (and more generally, why one would feel it at *any* time).

Reasons for non-agential pride are facts grounded in relatively stable properties belonging to the agent. Being thus-and-so (e.g. Indian) is not something that transpires in the way that achieving something does. For this reason, reasons for non-agential pride do

---

[23] "Ordinarily" is meant to allow for the possibility that, under the right circumstances, features like these may be objects of non-agential pride.

not expire in the way that agential reasons for pride might. This is not to say that facts about achievements are mutable. But, the bearing of particular achievements on one's well-being does change. When I learned to ride a bicycle, I was proud about being able to ride a bike. It was an achievement. I still happily ride my bike, but I am no longer proud about being able (or having learned) to do so. The time for pride has passed, plausibly because I do not (and presumably should not) view having learned to ride a bike as valuable to my life aims in the same way I did as a child. There are several questions raised by this observation, e.g. has my threshold for achievement grown such that my "learning to bike" no longer *is* an achievement for adult me?[24] Or, is it simply that my current attitudes are more sensitive to (or perhaps, biased towards) *recent* achievements? These are questions for another paper. Suffice it say that the 'motivating' reason for my bike-riding-pride has expired (even if my normative one has not). This difference in the temporal profiles of pride's (motivating) reasons renders it explicable why it is more likely for one to remain proud throughout one's lifetime of one's nationality, than one is about learning to ride a bike. Of course, some achievement may well become part of one's practical identity. This is no objection to the view on offer. In these fortunate cases, one's achievement acquires the temporal profile of a (motivating) reason for non-agential pride *because* this is the psychological upshot of representing some property as belonging to one's true self or practical identity.

Once it is acknowledged that non-agential reasons for pride are based in relatively stable properties, there isn't really a puzzle about why, being generally disposed to feel proud about being x (i.e., having some valued feature), I feel proud about being x *now*

---

[24] If "achievement" is construed as a thick-evaluative term, something like this might be defensible.

rather than two hours ago. Episodes of non-agential pride (like episodes of non-agential shame) can be occasioned in a variety of ways. When musician Ali Ward, speaking on behalf of her group (Hillfolk Noir), says "we are damn proud to be from Boise," the group's pride in coming from Boise, ID is an example of non-agential pride.[25] Now, if the emotion pride involves self-directed joy, as I claim it does, is this joy something that 'happens' in the case of non-agential pride? It is. While true that one's self-esteem can be heightened *generally*—e.g., after having his braces removed, Tim's self-esteem might *persistently* sit a notch higher—but I do not think this is how pride, the emotion, functions. That is, although I do not assume that pride is episodic in the same way as a headache, which unlike pride lacks propositional content, I do hold that pride has an episodic character. The important point here is not that pride is paradigmatically a positively valenced emotion— it is—but that agential pride and non-agential pride do not differ in that only one has a qualitative feel (notwithstanding this, the reasons underlying non-agential pride might give rise to a wider range of emotions).[26]

---

[25]   http://webcache.googleusercontent.com/search?q=cache:u-0O6j9_I-EJ:www.hillfolknoir.com/blog/+&cd=6&hl=en&ct=clnk&gl=us

[26] The claim that pride is episodic comes apart from the claim that it is essentially positively valenced. Davidson, for instance thinks that pride is constituted by formation of a *judgment*, and since judging is something we *do,* emotions have an episodic character. He departs here from Hume: "Hume surely did often, and characteristically, assert that a pleasant feeling, or a feeling of pleasure of a certain sort, was essential to pride, whereas no such feeling is essential; and more important, such an element does not help in analyzing and attitude of approval, or a judgment." Davidson (1980: 287). I am not concerned to argue that an affective core is *essential* to pride. My own view is that pride is paradigmatically valenced, though not 'accompanied' by valence. Feeling an emotion, I take it, is a contentful way of thinking, and the "positivity" of pride derives not from judgments about value but from the pleasure that, in adults at least, is ordinarily sensitive to states of affairs that one judges (inter alia) valuable. I have not argued for this view, and its postulation raises questions about the possibility of, say, the severely depressed person's feeling pride. That such examples arise as (potential) '*borderline* cases', however, points to the centrally of positive affect to pride.

Returning now to Ali of Hillfolk Noir, we can suppose that given her beliefs and values, she has a stable reason for non-agential pride about being from Boise. Thus she is disposed to see (or 'be motivated by') her enduring reason for pride when certain conditions are met, and about those cases we can say that her 'normative reason'—the set of considerations that are enduringly available to *justify* her attitude—generate a corresponding 'motivating reason', one attributable to her when she feels pride. Once this is granted, any number of things might direct her attention to this reason. Among Ali's reasons for Boise-pride might be the fact that Boise in home to *Built to Spill*, a pivotal indie-rock group that has produced excellent albums since the early 90's.

An example: consider the disappointment Ali Ward might feel if she were to realize that, contrary to what she had believed, she and her band were not really from Boise, ID. It is improbable for a person to be confused in this way about her origins, but we can imagine instilling in her false beliefs about being from Boise, along with pro-attitudes about Boise (of the sort she in fact has). With these attitudes in place, she might 'feel proud to be from Boise' (in spite failing to in fact be from Boise). Now, when she learns the truth, given the value she had imbued on the property 'being from Boise', it will come as a disappointment to her that she is not in fact from Boise. The epistemic conditions necessary for her reason for pride will be undermined, and although she might temporarily continue to feel for some joy when Boise comes to mind, she will also see that her reason for Boise pride has been extinguished, against her hopes.

Any adequate account of normative hope will need to be qualified in a way that renders intelligible how one might feel pride in spite of recognizing that one's attainment of the valuable goal is partly a result of luck. Two examples will be useful here. In 2006

Kiran Desai won the Man Booker prize for her novel, *The Inheritance of Loss*. It is not implausible that Desai saw winning the Booker prize as a valuable achievement for which she is responsible, and in so viewing the achievement, felt an elevation of self, relative to an ideal of e.g., good writing. That is, she felt proud. Additionally, if pride is ever appropriate for literary achievement, it will presumably be appropriate for Desai to feel proud about *this* achievement. In fact, speaking of the book, the judge for that year's prize said "[i]t is a magnificent novel of humane breadth and wisdom, comic tenderness and powerful political acuteness…Her mother will be proud of her."[27] (Kiran Desai's mother, Anita Desai, was shortlisted for the prize several times.) Desai's pride about winning this prize, a testament to her dedication to powerful writing, is presumably compatible with her appreciating the fact that there were, even if only a handful, other authors whose novels would have equally merited the laudation reserved for her. Of course, if Desai thought the selection process was *arbitrary*, she might by happy to receive the prize money, but it is difficult to imagine that she would feel proud. Thus, the pride she feels about her achievement is compatible with her recognizing that some luck was involved. So, even the normative hope of *agential* pride is capable of countenancing the importance of non-agential factors.

Similarly, two equally loving sets of parents might differ markedly in the pride they can take in their children. Good parenting (I don't pretend to know what this is) is conducive to but surely does not *guarantee* that later in life, one's child is an apt object of parental pride. Social influence and genetics, among many other factors, may intervene in such a way that renders one's child an object of more or less pride, and in some cases an

---

[27] https://www.theguardian.com/uk/2006/oct/11/books.bookerprize2006

object of shame. This point is complicated by the fact that the parental love generally underlies parental pride, such that the latter is not maximally sensitive to merit.[28] A parent whose positive attitudes towards his child were *soley* sensitive to merit (in either the agential or non-agential senses) is apt to be rather brutal, approximating something closer to a college admissions committee member or contest judge, than a parent.

It is tempting to think that pride essentially involve comparative evaluation, of the sort that the status-concerned parents might make about the relative successes of their children. Hume, for instance, thought that intragroup comparison was essential to pride.[29] In my view, however, intragroup comparison is an occasionally attendant but non-essential feature of pride. When a standard of excellence is socially acknowledged and aspired toward by many familiar parties, comparisons are unavoidable. Sometimes these comparisons serve a positive function. That is, they need to generate what Roberts (2013: 144) calls "invidious pride".[30] For, we speak of healthy competition in the athletic realm. Presumably, then, there is nothing objectionable in the comparative component of the gold-medal archer's pride. Now, any time one's pride is in part about one's *rank*, the content of pride will have an intragroup comparative feature. When the unit of comparison is social status, intragroup comparison is apt to be shallow. But, it would be too hasty to write off

---

[28] Even when merit is recognized to be high (and so, affectionate abundant) a parent might want her love to be rooted in something non-merit-sensitive. "Hal Incandenza for a long time identified himself as a lexical prodigy who—though Avril had taken pains to let all three of her children know that her nonjudgmental love and pride depended in no way on achievement or performance or potential talent—had made his mother proud, plus a really good tennis player." Foster Wallace (2006: 36)

[29] Hume, David (1967) [1738], 291

[30] The content of which is "it is important for me to have the personal worth that is established by my being or appearing to be equal or superior to R in respect X; I have that personal worth in some abundance because I am or appear to be superior to R in respect X" (Roberts 2013: 144).

all intragroup comparative pride (even outside the athletic realm) as motivated by concern for such status. Given that standards of excellence are often socially calibrated, one's feeling comparative pride can reflect the respect one has for the excellence of her peers. It is for this reason that the esteem-elevating objects of comparative pride will rarely be limited to one's rank (but include goods that are viewed as non-relative).

An example may help. On the *About* page of the Tufts Philosophy website one reads, among other sentences, the following: "Our master's degree program is world-renowned and is perennially rated the leading stepping-stone graduate program in the United States by The Philosophical Gourmet Report." [31] *Suppose* Tufts' philosophy department is proud of this fact (although I speak as if I can help myself to the idea of group-pride (e.g., a *department's* pride), feel free to change the example such that the pride is felt by one person, say, the department chair).[32] For this fact to be a source of pride, however, at least two background assumptions will presumably be present. The department will presumably hold that: a) other American terminal masters programs are valuable, and b) that their top-rating reflects the attitudes of those *competent to assess the merits* of terminal master's programs in philosophy. If a) did not hold it is difficult to see how being top-rated could be a source of pride. I do not deny that it's *possible* to care about one's rank independently of caring about the value of one's 'competition', but it is close to being incoherent. The very fact that the individual feeling comparative pride *restricts* the comparison-class suggests (or rather, is *revelatory of the fact*) that the non-relative qualities of the "competitors" are highly relevant. This is as clear in academic as it is in athletic

---

[31] http://as.tufts.edu/philosophy/about

[32] I do not deny that such statements may be motivated by considerations having nothing to do with pride.

evaluation. Even if *every* 10-year-old in the world took part in a competitive race against Olympic sprinter Usain Bolt, I wager Bolt would not take pride in finishing first. While b) is not obviously applicable in all cases of athletic ranking, in considering what counts as "good philosophy", philosophers rely on the estimation of the other philosophers; it is not, in other words, settled *a priori*, and so, a department may presumably learn something non-trivial about itself in coming to know, where its peers think it stands. So, while pride can be felt without intragroup comparison— consider pride for personal (especially idiosyncratic) achievement—many cases of the attitude may involve such comparison, and this does not by itself render heteronomous such pride.

Whether it is about an action or simply about some way one *is,* pride joyfully represents something valuable as reflecting positively on oneself. But, what is it to view some action or property as valuable in this context? The 'perceived value' component of pride *cannot* be understood in terms of one's making good on some normative expectation (or demand). This is just as true of agential pride as it is of non-agential pride. Banting may have felt proud about discovering insulin even if he never thought doing so was normatively required of him. By contrast, in blaming himself for stealing pears, Aquinas represents the norm proscribing theft as binding. I do not deny that one might feel proud about actions, the omission of which would generate in that person guilt. For instance, a lifeguard might feel proud about rescuing a drowning swimmer, and he would likely feel guilty (certainly, we would blame him) if he failed to do so, e.g., having fallen asleep on the job. But, the fact that normative demands underwrite blame of him does *not* mean that his pride consists in viewing himself to have fulfilled a normative demand—he fulfills these on uneventful days too. We do not feel proud every time we fail to violate (something

we take to be) a normative demand, nor should we. For this reason, we should construe the positive evaluation of emotions like pride in terms not of their objects' meeting a normative requirement, but of instantiating a normative ideal.

This is point is only somewhat appreciated by Taylor (1985: 43) who says that "[t]he person who is proud of some possession sees owning that thing as beyond his norm of expectations." Or again, according to Taylor (1985: 41) when one feels proud about something "she values it in that she sees it as a beyond her norm of expectation in the sense of its being better than for some reason or another she is, or others are, entitled to expect." A thought that follows naturally from this one is probably true, namely that persons with unattainably high normative expectations are unlikely to feel much pride. Still, we should not be misled into thinking that pride can be understood by reference to normative expectation. Something's being 'beyond one's normative expectation' is not by itself an informative notion, and our lifeguard, anyway, will in some cases feel proud about doing only what his job (and presumably, morality) demands. Similarly, suitably tempted to break a promise, Nina might feel proud about keeping it. This does not entail that she sees keeping her promise as "being better than for some reason or another she is, or others are, entitled to expect." Rather, I maintain, she feels proud because she sees keeping the promise as expressive of an ideal constitutive of the relationship to the promissee.

This view of pride as responsive to the exemplification of normative ideals bears some similarity to the view of praise-manifesting attitudes introduced by Martin (2013). For Martin, blame attitudes are expressive of, and respond to the violation of normative expectations (as per Wallace 1994), while praise attitudes are expressive of, and respond to compliance with, normative hopes. For Martin, however, normative hope *is essentially*

agential; "normatively hoping" is always normatively hoping *that some person acts* in one way or other. This point I think is unfounded. Normative hope *is* always something invested in persons. This I think is right, and so I accept the distinction Martin (2013) draws between non-normative and normative hope.[33] But a distinction should be drawn among kinds of normative hope. If my analysis of pride is right, normative hope will admit of two varieties: agential *and* non-agential. Martin's (2013: 140) general characterization of normative hope captures only the hope proper to *agential* normative pride. She writes, "[w]hen we invest hope in people, we stand ready to feel and justify feeling disappointment in the event that they fail to act as we hope and gratitude when they succeed…to invest hope in a person is to relate some norm to her aspirationally, to hold it up to her as something with which she has decisive reason to comply". When we invest hope in persons (including ourselves) *non-agentially*, we relate some norm to that person (including oneself) aspirationally, but in such a way that 'acting on decisive reasons' is not at issue. A view of normative hope that ignores this is apt to restrict the bounds of one's practical identity to the boundaries of one's agency. But we are more, normatively speaking, then what we do.

---

[33] Where, "[*n*]on-normative hope [which] involves a stance taken toward the probability one assigns to the hoped-for outcome, such that it licenses certain feelings and not others" (Martin 2013: 130).

3

Responsibility as Estimability

0. Introduction

Where philosophers speak of moral responsibility, they often also speak of "accountability". Following Gary Watson (1996), it is customary for theorists of responsibility to distinguish between two types of responsibility: attributability and accountability. For an agent to be responsible for an action in the attributability sense is for that action to be expressive of a privileged set of that agent's attitudes, such as her evaluative commitments or cares.[1] In contrast to actions performed on a whim, under hypnosis, impulsively, under situational constraints, etc., some actions intuitively disclose

---

[1] There is debate over what features of the agent's psyche ground attributability. On Platonic views, this role is played by the agent's judgments about what ends are valuable (Watson 1987), or *most* valuable (Watson 1975). On Humean views, by contrast, actions are attributable to the deep self if they issue from a desire that the agent desires to be efficacious (Frankfurt 1971), or if they reflect what she cares about, where caring about x can stand in tension with, if not plainly conflict with, one's judgments regarding x's value (Frankfurt 1982). Hybrid views are also available; on the ecumenical view proposed by Shoemaker (2013; 2015), both evaluative commitments and non-cognitive cares may express one's deep self.

what one *stands for*, or who *the agent really is*— what is sometimes referred to as one's "practical identity" or "real self" (Watson 1975; 1996; Wolf 1987; 1990). These kinds of actions are attributable to us in the sense that they reveal the quality of our characters, and importantly, this includes excellences and faults of character. For an agent to be responsible in the more robust *accountability* sense, by contrast, is for her to be an apt target of a range of interpersonal responses, such as those of (especially) affective blame and praise, or what P.F. Strawson (1962) calls the "reactive attitudes". An agent's being responsible in the accountability sense includes his being generally *held accountable* by other agents, where this is a matter of being bound by mutually recognized normative expectations, demands, or requirements (typically understood to track moral obligations or directed duties). Depending on the quality of one's actual conduct—particularly on the quality of will or *regard* for others that the conduct manifests— an accountable agent may on some instance be *held to account*. That is, if the agent flouts some normative requirement, then depending on factors like the severity of the violation, the wrongdoer's interests are apt to be affected by other accountable agents—especially the person he *wronged*— e.g. by being treated with attitudes of blame (like resentment and indignation), and through modifications in the blamer's trust and good will towards the wrongdoer. It is for these reasons that accountability raises questions regarding fairness and desert that attributability does not.

It is not difficult to see how responsibility in the attributability sense encompasses not only the negative pole of responsibility, but the positive one too— i.e. not only blame, but also praise. If one's telling the truth in some challenging situation is expressive of her honesty (rather than, say, fear of being caught in a lie), that action is naturally construed as expressive of an excellence of the truth-telling agent's character. A praiseworthy action is

in this sense attributable to her.[2] With accountability, however, the connection to praise is less clear. This is because "normative demand" is the concept focal to accountability. According to the accountability theorist, in blaming an agent I *hold him to a demand* (in a sense to be examined below) for e.g. proper regard or respect, which—given the appearance of transgression— calls for the supposed wrongdoer to offer explanation, justification, excuse, apology, etc. In praising you, by contrast, I am not in any obvious sense holding you to a normative demand. While the persons we praise will normally *in fact* be bound by normative demands (e.g. the obligation not to harm us for fun), this does not mean that one's praise is *itself* a way of holding an agent to a normative demand, nor is it clear what it could mean for praise to issue a demand. Unlike attributability, then, accountability appears tailor-made— or at least, asymmetrically suitable— for elucidating the negative pole of moral responsibility, namely that concerned with interpersonal blame.

In spite of this peculiarity, responsibility theorists standardly take normative demand to be the focal concept not only for interpersonal blameworthiness, but for a full-scale type of responsibility, one that includes "accountability praise". This assumption

---

[2] Attributability in this agentially relevant sense should be distinguished from a thinner sense in which we might say that an action is attributable to some agent (rather than to another) just in case it can be traced to *some* mental state of hers. This point is made by Shoemaker (2015: 60 n.25). Note that not everyone accepts that attributability *is* a genuine form of responsibility. See Levy (2005) for the challenge that attributability theorists illicitly collapse the distinction between an act's being bad and the agent's being blameworthy for it. Though I refer to in this paper to attributably responsibility (and responsibility in the 'more-than-attributability' sense), neither the negative nor positive arguments of this paper are affected by skepticism about the view that attributability is a form of responsibility. Given the widespread acceptance of the Watsonian distinction between attributability and accountability (considered as forms of responsibility), however, I put this skeptical challenge aside for the remainder of this paper.

embodies what I call the "demand prejudice" in the responsibility literature.[3] The demand prejudice is the assumption that the concept of normative demand (or its any close cousins, e.g. normative requirement, expectation, requirement) is the organizing concept for an overall theory of moral responsibility, i.e. for a theory of responsibility that includes a view of both blame *and praise* (see e.g. Strawson 1962; Darwall 2006; McKenna 2012; Nelkin 2016). Plausibly a symptom of a broader neglect of praise and praiseworthiness in the responsibility literature, the demand prejudice is prevalent among theorists of moral responsibility, especially among proponents of accountability theories of responsibility. One expression of the demand prejudice is particularly pertinent here, namely the assumption that our responsibility practices include a form of interpersonal praise that corresponds to what might be called "accountability-praise", or praise in the accountability sense.[4] Michael McKenna gives clear expression this way of conceiving of responsibility when he writes:

> A *morally responsible agent,* in the sense of moral responsibility I wish to consider, is *accountable* for her conduct. She is one who can be *held* morally responsible for what she has done, can be an appropriate target of praise or blame, and of relevant moral emotions, such as gratitude or resentment, as well as moral approbation or indignation.[5]
>
> (McKenna 2012: 11-12)

---

[3] Since I do not establish until section 2 the prejudicial quality of what I call the "demand *prejudice*", the reader may— in the interim—substitute "demand *assumption*" for this label.

[4] Unless otherwise noted, it is this assumption that I will have in mind in referring to the *demand prejudice*.

[5] McKenna and Coates (2015) too presuppose 'accountability praise' in writing that "[a] person who is a *morally responsible agent* is not merely a person who is able to *do* moral right or wrong. Beyond this, she is *accountable* for her morally significant conduct. Hence, she is, when fitting, an apt target of moral praise or blame, as well as reward or punishment."

McKenna here takes responsibility in the accountability sense to encompass not only our responses of interpersonal blame, but also those of interpersonal praise.

The picture of responsibility painted by the accountability theorist is one in which responsible agents stand ready to exercise a kind of normative *authority* over one another, an authority to confront another with demands, to 'hold each other *to'* requirements, and to 'hold wrongs *against'* another when a normative requirement is flouted (Darwall 2006: 12; Nelkin 2016: 180-1). This nexus of concepts has the virtue of making clear what is *interpersonal* about accountability, but it does so at the cost of obscuring the nature of interpersonal praise. Consider David Shoemaker's (2015: 87) gloss on accountability, where "to be accountable for something is to be liable to be appropriately *held* to account for it, which is to be eligible for a range fitting responsibility responses with a built-in confrontational element". The idea that 'to be *praiseworthy* for x is to be *liable* to be held to account for x' is awkward, at best. This much is acknowledged by Watson himself, though in spirit of optimism, when he writes that "[w]e seem to have a richer vocabulary of blame than of praise…To be held liable is to be on the hook, and we lack a ready phrase for a positive counterpart to the "hook". But clearly we do have a counterpart notion" (1996: 283; 284). Optimism about the prospects of "praise in the accountability sense" has endured among accountability theorists, but so has the attendant awkwardness. Echoing Watson, Dana Nelkin recently writes:

> Is there a praising "stance" that corresponds to the blaming one of "holding an offense against one"? We do not have an expression that directly parallels this one, such as "holding a good deed *to* one." And it seems that even if there were a concept for which such an expression would be apt (as I am inclined to believe there is), it must be acknowledge that it would not be the mirror image of that for blame in all details.   (Nelkin 2016: 186)

Given, however, that demand is the "central notion" for accountability theorists—and Nelkin (2016: 180-1) is no exception— we should be skeptical about the prospects of rendering intelligible "praiseworthiness in the accountability sense" (Nelkin 2016: 186), and more generally, about the viability of a picture of interpersonal praise that differs from accountability blame only in its *details*.

The criticism that the concepts proper to accountability are ill-suited for elucidating praise is not a novel one. It is raised, in one way or another, by each of Bennett (1980); Eshleman (2014); King (2014); Martin (2014: 118-125); and Russell (2013). Why, then, does the demand prejudice persist? The part of the answer that concerns me is that a satisfactory counterpart to accountability has yet to be outlined.[6] Responsibility in what I call the "estimability" sense is poised to fill this gap. Praise in the more-than-attributability sense, I propose, should be understood in terms of one's stance towards agents of 'holding in esteem'. Like the accountability, estimability is an essentially interpersonal form of responsibility that invokes questions of desert and fairness.[7] But unlike accountability, estimability is not structured around the deontic concepts of duty, demand, obligation, expectation, and the like. Rather, estimability is organized around axiological (or evaluative) concepts, ones that reflect the *ideals* that we hold others (and ourselves) *up to*.[8]

---

[6] This is not to say that the abovementioned figures have *failed* to provide a suitable counterpart to accountability. King and Eshleman notably do try, but they think praiseworthiness is to be understood solely in terms of attributability-responsibility, an idea a take issue with below.

[7] A common way of articulating blame's warrant is in terms of *desert,* but depending on how rich a notion of 'desert' one has in mind, accountability theorists may appeal to normative considerations other than (basic or non-basic) desert to account for the warrant of holding responsible. I return to this issue in section 4.

[8] A note on the contrast between the deontic and the axiological (/evaluative): the normative domain can be divided into two: i) the axiological or evaluative (from *axos* and *valores*, both meaning 'possessing worth' in Greek and Latin, respectively), the core concepts of

Estimability, I hasten to add, is a form of responsibility already recognized in our practical lives. That is, we regularly hope that others will aspire towards and satisfy various ideals without always (or *thereby*) requiring or demanding that they do so. What is lacking is the philosophical vocabulary of praise requisite for theorizing this sphere of responsible agency. This paper is an effort to chart out such a vocabulary. But first, more must be said to establish the need for this undertaking, i.e. to establish the erroneousness of the demand prejudice and its stronghold within accountability views.

In section 1, I provide an overview of the rise to prominence of the concept of "accountability". In section 2, I show why the attendant demand prejudice is a prejudice, i.e. why the concepts proper to accountability (normative demand, directed duty, claim-right, etc.) are incapable of providing us with an understanding of either the content of praise attitudes, or a way of understanding what is characteristic of our responses to agents we take to be praiseworthy. To fill the gap left by the previous section, I advance in section 3 a novel conception of responsibility, which I call responsibility in the "estimability"

which are "good"/"valuable" and "bad"/ "disvaluable", i.e. 'thin' evaluative concepts implied, and rendered more determinate, in a vast range of "thick" (and descriptively rich) evaluative concepts, like "generous", "admirable", "courageous", "lewd", "sinister", "devious", etc.; and ii) the deontic (from *deon*, that which is binding, in Greek), the central concepts of which are "right" and "wrong", along with "obligatory", "ought", "duty", "permissible", "impermissible", etc. It is worth noting that "normative" is *sometimes* used as synonymous with "deontic", perhaps in part because the descriptive/normative gap is often spelled out in terms of is/*ought* (and "ought", at least in one primary sense, is a deontic concept). In any case, I take the deontic and evaluative/axiological to be species of the normative, in a manner reflected by e.g. Dancy's (2000: 29) claim that, "[w]ithin the normative realm, considered as a whole, there are two potentially distinct parts. There is the evaluative, and there is the deontic. The evaluative is the realm of value, of good and bad, of evil and benign. The deontic is the realm of 'oughts'— of requirements, of demands, of right and wrong". I take no stance in this paper on the meta-normative question of which of the two is conceptually prior (if either— see Wedgewood 2009). Lastly, I speak primarily of deontic and axiological *concepts* (rather than properties, say) because my focus is on the conceptual adequacy of "normative demand" for understanding praise.

sense. In section 4 I explore the prospects for a general account of interpersonal responsibility that may include accountability and estimability as species.

1. Sources of Accountability

The idea that accountability constitutes a distinctive type of moral responsibility is introduced by Watson (1996) and Wallace (1994), both finding inspiration in Strawson's (1962) "Freedom and Resentment". After briefly reviewing the demand prejudice's roots in Strawson, I turn towards its subsequent developments with the hope that doing so will position us to see the accidental nature of identifying "accountability" with a type of responsibility (i.e. blameworthiness *and* praiseworthiness).

1.1   Strawsonian Demands

A significant feature of our everyday interpersonal lives, Strawson (2003 [1962]: 78) claims, is the "demand for goodwill or regard which is reflected in our ordinary reactive attitudes," where reactive attitudes are emotions (like resentment, indignation, gratitude, and admiration) that are constitutive of praise and blame. Rejecting the idea that moral responsibility's justification is hostage to metaphysical issues regarding the compatibility of free will with determinism, Strawson proposes that we hold onto our pre-theoretical commitment to praising and blaming agents by attending to the observation that moral responsibility is part and parcel of a kind of interpersonal engagement to which we are naturally disposed and as such stands in no need of justification from without. We do recognize circumstances in which agents are excused or exempt from particular tokens of our reactive attitudes. But those are circumstances in which some feature of the agent's

capacities or state (e.g. some psychological abnormality) explains away the appearance that he has flouted a demand for proper regard. Those are circumstances, in other words, where the reactive attitudes are undermined by standards that are *internal* to our responsibility practices.[9]

I will not assess Strawson's argument for the view that questions of moral responsibility are insulated from considerations external to those practices (e.g. general facts about determinism). The aim here is rather to locate in Strawson the seed for the view of accountability developed by Watson and others. It is not difficult to find various claims to the effect that the reactive attitudes are constitutively linked to *demands* and *expectations*. Indeed, the starting point for Strawson's positive proposal is reflection on the "kind of importance we attach to the attitudes and intentions towards us of those who stand in [interpersonal] relationships] to us, and the kinds of reactive attitudes and feelings to which we ourselves are prone. In general, we <u>demand</u> some degree of goodwill or regard on the part of those who stand in these relationships to us" (Strawson 2003 [1962]: 76). It is an inextricable feature of human life as we know it that our relationships are conditioned by the general demand for others' good will. Our tendency to undergo reactive attitudes (those of blame and praise) is itself an expression (or perhaps a constitutive element) of this general human demand. To illustrate that this appeal to demands is not a 'one-off' for Strawson, here are several more examples:

---

[9] Strawson is sometimes taken to hold the view that 'because' it is psychologically necessary for us to feel token reactive attitudes, external criteria cannot render unjustified instances of these attitudes (token-naturalism). Following Russell (1992) I've presented Strawson only as espousing "type-naturalism", the more modest view on which it is the disposition towards reactive attitudes that requires no external justification.

> The personal reactive attitudes rest on, and reflect, an <u>expectation of</u>, and <u>demand for</u>, the manifestation of a certain degree of goodwill or regard on the part of other human beings towards ourselves; or at least on the <u>expectation of</u>, and <u>demand for</u>, an absence of the manifestation of active ill will or indifferent disregard.
>   (Strawson 2003 [1962]: 84)

> [Reactive] attitudes involve, or express, a certain sort of <u>demand for</u> inter-personal regard.   (Strawson 2003 [1962]: 85)

> The <u>making of the demand</u> *is* the proneness to such [reactive] attitudes…The partial withdrawal of goodwill which *these* [in contrast to non-reactive] attitudes entail, the modification *they* entail of the <u>general demand</u> that another   should,   if possible, be spared suffering, is, rather the consequence of continuing to   view him as a member of the moral community; only as one who has offended against its <u>demands</u>.   (Strawson 2003 [1962]: 90)

What kind of thing does Strawson take the relevant sort of demand to be? We know that it is to be "qualified as 'moral'" such that *moral demand* is what is meant; the demands in question are those of persons, but they are also described as the *moral community's* demands (2003 [1962]: 90). Strawson also indicates that the demands in question those reflective of moral obligation (2003 [1962]: 86). Further light is shed on Strawson's "demands" in a slightly earlier paper, "Social Morality and Individual Ideal" (1961), where Strawson advances a "minimal social interpretation of morality". There the demands are further qualified as "socially sanctioned":

> The fundamental idea is that of a socially sanctioned demand made on an individual in virtue merely of his membership of the society in question, or in virtue of a particular position which he occupies within it or a particular relation in which he stands to other members of it.   (Strawson 1961: 105)

For the socially sanctioned demands to be reflective of moral obligations (and not, say, the orders of an oppressive ruling-class), the community must be one in which there is a "reciprocal acknowledgment of rights and duties." (Strawson 1961: 107) The community

must, in other words, be one in which agents hold one another *mutually accountable*. For, "if I have no foothold at all in the sanctioning part of society, and if no interest of mine is safeguarded by the system of demands to which I am subject, then, in fulfilling a demand made upon me, I may indeed, in one sense, be doing what I am obliged to do; but scarcely what I am *morally* obliged to do." (Strawson 1961:108) Our proneness to reactive attitudes, then, can be attributed to the reciprocal (obligation-reflecting) demands to which we, as members of a moral community, hold one another, where being 'held' to such demands is a matter of being liable to the (social and emotional) sanctions of others.[10]

While the above passages do not provide a theory of responsibility in terms of 'accountability',[11] the idea that moral responsibility is a matter of holding agents to interpersonal expectations and demands that are reflective of moral obligations—*this idea* is central to Strawson, and it understandably serves as the basis for the broadly Strawsonian picture of accountability developed by Watson and others. Important differences exist between the views in question, but we can focus on what they have in common.

1.2 Accountability after Strawson

[10] —"in virtue merely of [their] membership of the society in question, or in virtue of a particular position which [they] occupies within it or a particular relation in which [they] stand[] to other members of it." (Strawson 1961: 105) The language of 'sanctions' is missing from "Freedom and Resentment", but that essay's reference to the 'infliction of suffering' and 'withdrawal of good will' plausibly comes to the same thing.

[11] The term 'accountability' is itself absent from "Freedom and Resentment". It is therefore unobvious why Bennett's [1980] influential reply to Strawson is centered on the concepts of, and conditions for, accountability: "By 'accountable' I shall mean 'blameworthy or praiseworthy': someone is 'accountable' for an action, in my usage, if a blame- or praise-related response to the action would not be inappropriate. And my concern is with a problem about the conditions for 'accountability' in this sense" (Bennett 2008 [1980]: 49). Like Strawson, Schlick (1939) fails to treat "accountability" as technical term, though Bennett attributes to him a 'theory of accountability' (Bennett 1980 [2008]: 51).

Watson introduces "accountability" to refer to the interpersonal form of moral responsibility that Strawson outlines by reference to the reactive attitudes. The centrality of the concept of "normative demand" is explicit on Watson's account of accountability, on which "holding responsible is a three-term relationship in which one individual or group is held by another to certain expectations or demands or requirements" (Watson 2004 [1996]: 274). Watson's proposal is that moral agents bear a special class of expectations of one another, and that these expectations reflect requirements (or duties) that ground corresponding demands. Unlike the attributability responsibility that an agent has for an action in virtue of its being expressive of her evaluative commitments or cares, an agent's being *accountable* for some action warrants (at least pro tanto) certain kinds of *responses to,* and sometimes, *treatment of,* her. More concretely, "[h]olding accountable…involves the liability to sanctions. To be entitled to make demands, then, is to be entitled to impose conditions of liability," for, "blaming responses (at least potentially) affect the interests of their objects adversely" (Watson 2004 [1996]: 275; 280). For this reason, accountability raises questions regarding *fairness* that attributability—otherwise referred to as the "*aretaic* face" of responsibility—does not (Watson 2004 [1996]: 273).[12]

What is it to be *fairly* held accountable? In Watson's view, this requires that the blamee be capable of understanding the relevant normative expectations.[13] More precisely, the intelligibility of an agent's being subject to a demand depends on the agent's capacity to understand the relevant normative demands as reason-giving. Demands have

---

[12]"Because many of these appraisals concern the agent's excellences and faults—or virtues and vices—as manifested in thought and action, I shall say that such judgments are made from the *aretaic perspective*" (Watson 2004 [1996]: 266).

[13] And further, though less importantly for my immediate purposes, to be able to control one's conduct in light of them.

constitutive *uptake-conditions*. When these conditions are not met, the agent cannot meaningfully be construed as being under a normative demand, and to treat him as such is (at least potentially) to treat him unfairly. The demand must, as it were, be heard (or at least be *audible*) in order for it to have the status as a normative demand on the agent's conduct. This communicative approach to responsibility is traceable to an earlier, and comparably influential, paper in which Watson writes:

> The negative reactive attitudes express a *moral* demand, a demand for reasonable regard…Since the negative reactive attitudes involve this demand, they are not (as fully) appropriately directed to those who do not fully grasp the terms of the demand. To be intelligible, demanding presumes understanding on the part of the object of the demand. The reactive attitudes are incipiently forms of communication, which make sense only on the assumption that the other can comprehend the message. (Watson 2004 [1987]: 229-230)

The idea that the reactive attitudes are "incipient forms of communication" or "forms of *moral address*" (Watson 2004 [1987]: 230-1) arguably adds further content to Strawson's own picture, and has been developed by 'Strawsonians' in several ways. Some theorists maintain that praise and blame are *paradigmatically* overtly communicated, and for that reason overt blame and praise are explanatorily prior to their private counterparts (McKenna 2012). On this type of view, the norms of blame and praise are norms governing expressed blame and praise. Others by contrast hold that the reactive attitudes, even when considered as unexpressed or private emotions, are "communicative *entities*" in virtue of having the function of evoking uptake of their representational contents, which contents are norm-governed independently of their possible expression (Macnamara 2013; 2015a). Differences among communicative views of responsibility do not matter for our purposes. What matters is what they have in common: commitment the view that interpersonal moral

responsibility [14] is to be understood as a three-place relation between addressor and addressee, where a condition of eligibility for occupying either position is the ability to *understand* the form of moral address that blame (and ostensibly, praise) embodies.

In addition to adding substance to the idea that accountability is inherently interpersonal, the communicative approach provides a clear way of specifying the conditions under which agents may be exempt from accountability responses, namely the conditions under which they lack the capacity to understand and be guided by moral demands (Watson 2004 [1987]: 228-9). [15] The communicative approach to moral responsibility is emphasized to varying degrees by Strawsonians, but at a minimum, proponents of this approach are committed to understanding interpersonal responsibility (i.e. accountability) in terms of the mutual intelligibility and bindingness of a certain of a class of reasons that are manifested through (sometimes sanction-like) reactive attitudes as moral demands. That Jay Wallace thinks of accountability along these lines is evident in the following passage:

---

[14] Contrasted with attributability, which is not inherently in interpersonal.

[15] I put aside the difficult question of what the possession of such a capacity consists in, noting only that psychopaths are often posited as paradigmatic agents *lacking* the capacity. While not ubiquitous, Watson holds this view, and so does not think psychopaths should be held (fully) accountable, reasoning that "it makes no sense to address moral demands to [persons who cannot understand them, like psychopaths], as though these could be intelligible to them (or intelligible in the right kind of way) from their deliberative standpoints. It is partially in virtue of the competence embodied in the moral sentiments that it is possible for us to have relations of mutual accountability" (Watson 2011: 318). In denying that normative demands are intelligible to psychopaths *in the right way*, part of what Watson means is that psychopaths are incapable of seeing moral demands as issuing *non-instrumental* reasons for action (2011: 309). (See also Darwall 2006: 89; Fischer and Ravizza 1998" 79). On another kind of view, one that locates the psychopath's exemption in his *motivational* profile, the psychopath is exempt from our accountability practices in virtue of lacking the capacity to be non-instrumentally motivated by moral reasons, which itself is due to his lacking the capacity to care about the source of those reasons, namely moral agents (Shoemaker 2007: 84-5).

> the stance of holding people accountable involves a susceptibility to sanctioning behavior whose function is essentially to express the reactive emotions. Insofar as these responses are bound up with the moral obligations we accept, they will only be fully intelligible as forms of expression when addressed to people who are capable of grasping moral reasons.[16]   (Wallace 1994: 164)

The stance of holding accountable, also referred to as "holding someone to a demand", is for Wallace a *sui generis* way of relating to persons as moral agents (Wallace 1994: 45). The intelligibility of relating to a person in this way requires that she be capable of understanding, as non-instrumentally reason-giving, the relevant demands, which may find expression in blame-manifesting attitudes like resentment, indignation, and (in the case of holding oneself accountable) guilt. So like Watson, Wallace takes himself to be working within a Strawsonian framework in which the stance of holding persons responsible is characterized by a susceptibility to a range of emotions that express normative "demands" or expectations" where these are understood not only in the psychological sense, but as referring to "that which we demand or expect" of persons qua moral agents, such that normative demands and expectations are "equivalent ways of expressing the notion of a practical *requirement* or *prohibition*" (Wallace 1994: 22).

In addition to Watson, and Wallace, a communicative and demand-based view of responsibility is endorsed by each of Darwall, McKenna, and Shoemaker:

> reactive attitudes are "forms of communication" that are simply unintelligible in their own terms without the presupposition that their objects can understand what is being said and act on this understanding…what we presuppose in second-

---

[16] See also Wallace's (1994: 215) claim that "[t]he basic principle of fairness implied by the exemptions, I have suggested, is that it is reasonable to hold a person to a demand only if the person possesses the general power to grasp and comply with the reasons that support the demand."

personal address is *second-personal competence,* that those we address…can take a second-personal perspective on themselves and act on reasons they accept from that point of view (by making the relevant demands of themselves).

(Darwall 2006: 75)

Holding a person morally responsible…involves adopting a *practical* attitude toward the person by which those holding morally responsible are disposed to place certain moral demands upon her, have certain expectations of her, and praise and blame her depending on how she conducts herself… Exempted agents who are incapable of understanding the forms of moral communication put to them via our practices of holding morally responsible are incapable of *participation* in [these] practices.

(McKenna 2012: 22; 7)

[O]ur practices in voicing the praise and blame expressive of holding someone morally responsible, in the paradigm case, consist of an interplay between at least two agents, one who addresses a moral demand to the other via the praise or blame and the other who ostensibly hears, understands, and either accepts or rejects the demand…If one lacks the capacity for this sort of reason-based address and exchange—a specific type of normative competence—then one must be outside the boundaries of the moral community, beyond the reach of a certain type of relationship. (Shoemaker 2007: 70-1)

What has emerged so far is a picture of accountability-responsibility on which an agent's being accountable requires—both as a conceptual matter and a moral one—that that agent be capable of understanding a certain form of moral address, particularly, the *demands* constitutive of membership in the "moral community" of accountable agents.

Although it is not essential to a theory of accountability that normative demands be understood in terms of Hohfeldian claim rights— with corresponding directed duties or obligations—it is noteworthy that both Darwall and Wallace understand normative demands in these terms.[17] On this kind of view, A has a claim(-right) that B ϕ if and only

---

[17] "According to Hohfeld, someone has a claim right to another person's doing something only if that person has an obligation to her to do that thing. And this consists not simply in

if B has a duty to A to ϕ (where ϕ refers to type of act or omission). Given its relational structure, an analysis of normative demands in terms of claim-rights (and corresponding directed duties) is well suited for explicating accountability. Although talk of directed duties or obligations is absent in Watson, his view of accountability is broadly consistent with the Hohfeldian deontic picture outlined by Darwall and Wallace. This, at least, is strongly suggested by claims like, "holding responsible is a three-term relationship in which one individual or group is held by another to certain expectations or demands or requirements" (Watson 2004 [1996]: 274). There is bound to be disagreement among Wallace, Darwall, and Watson regarding how to delineate the relevant demands or requirements, but what matters for my purposes is their endorsement of the idea that one is a responsible agent in the robust, accountability, sense in virtue of one's sensitivity to a certain kind of "moral address", which is understood in terms of an agent's capacity to understand certain deontic considerations, like requirements and demands.

---

its being the case that the other ought or has good and sufficient reasons to do it…but in the claim-holder's authority to demand compliance and, perhaps, compensation for non-compliance. Rights are thus associated conceptually with second-personal reasons." (Darwall 2006: 18; 19; 273); Wallace ms. 2-3; cf. 26 n.2: "The idea of relational requirements in the most general sense is implicit in talk of moral and legal rights…More specifically, it is familiar from discussions of claim-rights, in the sense made familiar by Wesley Newcombe Hohfeld…[it strikes me as] promising to interpret morality in relational terms, as a set of requirements on agents that are like the obligations in the Hohfeldian domain in being constitutively connected to claims that other individuals have against us. While Wallace (1994) does not identify Hohfeld as a guiding light as Wallace (ms.) does, a relational and demand-based theory of accountability is common to both. See e.g., Wallace's claim that "[m]oral standards of right and wrong purport to define constraints on agency, but they also purport to provide a basis for interpersonal accountability relations between individuals, articulating what we can expect of each other as each of us pursues our private ends. Disregard of such interpersonal expectations by an agent thus has normative implications for other parties, giving them reason to adjust their attitudes and behavior in response, in the characteristic register of blame." Wallace ms. 103; Ch. 3 generally.

The question at the heart of my paper—or at least its negative part—is: why identify normative address with one of its varieties, namely, normative *demand*? In contrast to demand, normative address can take the form of invitation, affirmation, advice, request, etc. If what is distinctive of accountability is moral communication connected to interest-affecting responses, why have we limited ourselves to a *single kind* of normative address, i.e. demand? This is offered partly as an internal critique. In claiming that "holding responsible requires the intelligibility of moral address…[where] a condition of such address is that the other be seen as a potential moral interlocutor" Watson (2004 [1987]: 235) seems to avail himself of full range of forms of moral address. It is for this reason unfortunate that Watson effectively identifies moral address with moral demand, e.g. in seeking "to construe the exempting conditions as indications of the constraints on intelligible moral demand *or, put another way, of the constraints on moral address*." (Watson 2004 [1987]: 228-9, italics added). This identification is especially puzzling given that Watson acknowledges the existence of a kind of praise that is on par with accountability-blame. Though Watson's comments on the topic are mostly suggestive, they contain the insight that robust (i.e. *more than merely attributable*) praiseworthiness might be connected to the concepts of moral address and desert. I develop this suggestion below. But first, more needs to be said to establish why accountability won't do.

2. Holding Accountability to Account

Accountability won't do. That is, the deontic concepts proper to accountability are ill-suited for elucidating the nature and norms of praise. This is not an original idea. It is

present already in an early response to Strawson by Jonathan Bennett (1980: 42), who writes: "I doubt if 'demand' really covers all the ground: I can find no place for it in describing such undisappointed reactive feelings as those of gratitude and reciprocating love." This of course is not an *argument* against the idea that normative demand is the focal concept for such attitudes. But, neither Strawson nor any of his followers offer an argument for the conclusion that praise consists in holding other to demands. (Indeed, Wallace (1994: 25-33) offers on Strawson's behalf an argument *against* this conclusion.) What could such an argument look like? Recall that Strawson's methodology involves characterizing moral responsibility in terms of the dispositions proper to our interpersonal relationships, *as we find them*. The measure of such a characterization is intuitively its ability to reflect what we are doing when we, inter alia, praise and blame one another. While one kind of error-theorist might coherently offer a theory of praise on which normative demand is, despite all appearances, the focal concept, a strategy like this is unavailable to the Strawsonian accountability theorist, for whom "[o]nly by attending to this range of attitudes [like gratitude, resentment, etc.] can we recover from the facts as we know them a sense of what we mean, i.e. *all* we mean, when speaking the language of morals, we speak of desert, responsibility, guilt, condemnation, and justice." (2004 [1962]: 91). Can a demand-based account of responsibility live up to our experiences as agents who respond to the praiseworthy actions? Like Bennett, several philosophers have thought not. The intuitive objection is expressed clearly by Matt King:

> Accountability draws from our practices of holding each other responsible, which, unlike holding others to *be* responsible, commits us to a stance of expectation and demand, one that invites the imposition of sanctions…None of this looks remotely like praise…in praising the agent, I am not thereby holding her to any demand…If Sherry has donated money to Oxfam, Gina doesn't hold her accountable for it.

> There's nothing to 'account' for. Gina may, of course, hold Sherry accountable in general, for Sherry is a moral agent. But in praising her act of generosity, she is not thereby holding her to a demand to be generous.　　　　(King 2014: 420)

Andrew Eshleman too denies that the accountability model is capable of shedding light on our praise responses, for "the kinds of reasons presupposed by our praise-related attitudes are typically not the same kind, of reasons that undergird those involved in regarding someone as accountable to blame" (Eshleman 2014: 221). This general point, I argue in the next section, is right. Unfortunately, King and Eshleman take our shared intuition about the absence of demands in praise to speak in favor of the thesis that the province of praise is solely that of *attributability*-responsibility (Eshleman 2014: 235; King 2014: 420-1).[18] I think this is a mistake. Praise in the attributability sense is a genuine phenomenon, but so too is praise in a more robust, interpersonal, and desert-entailing sense that implicates fairness, namely praise in the estimability sense. My explication and substantiation of the claim will have to wait until the next section. For now, more must be said to show that 'accountability won't do'.

Consider the suggestion made by some accountability theorists that praise is conceptually connected to normative demand in that praise is the proper response to an agent's *exceeding* a normative demand (Darwall 2006: 73; McKenna 2012 7-8; 49). Call this the "supererogation strategy". The concept of the supererogatory makes essential reference to normative demands or duties, for a supererogatory act is a morally good act that, in some sense, involves doing *more* than is demanded by duty. Praise, it would seem, presupposes that we hold other agents accountable—that is, to normative demands— for

---

[18] Bennett is silent on this issue; his paper is largely negative (offered as it originally was, as commentary on "Freedom and Resentment").

such a stance is necessary in order to recognize that some agent has gone *beyond* what is normatively demanded of him. This is effectively what Darwall says when he claims that gratitude, which for the Strawsonian is a paradigmatic attitude of praise, is

> parasitic on legitimate claims or expectations. We are appropriately grateful when people benefit us or act as we wish when we lack any relevant claim or expectation of them… And gratitude's natural expression is also second-personal, a grateful addressing of the benefactor that reciprocally recognizes that he has benefited us beyond what we had any claim to expect.   (Darwall 2006: 73)

Can the supererogation strategy, then, provide the necessary link between normative demand and praise? If so, the accountability-theorist might make good on the idea that normative demand is the focal concept not only for a robust and interpersonal type of blame, but also for an analogous type of praise, namely *accountability* praise.

The supererogation strategy is unsuccessful for several reasons. The most glaring objection is that exceeding a normative demand is intuitively *unnecessary* to warrant praise. The person who keeps an important promise in face of the temptation to break it does no more than her duty, but we might praise her nonetheless. The same is true of the lifeguard who saves the (now grateful) parents' child at the pool he oversees. The lifeguard does only what is his duty. The parents, we can suppose, recognize this fact, as they would have blamed him (to say the least) had the lifeguard failed to save the child. The same is intuitively true about the promise-breaker. But, if this is so, such acts cannot be viewed by the person praising as supererogatory, for supererogatory acts are characterized in part by their *optionality*. For a supererogatory act to be 'beyond duty' is in part for its non-performance to be permissible, and so, not blameworthy. Since it is unnecessary that I take the target on my praise to have acted in excess of duty, praise cannot be understood as responsive to the supererogatory. While the accountability theorist might claim that the

consistent fulfillment of duty is supererogatory, even if this claim could be justified, it would be of little help given that *single acts* of duty-fulfillment are sometimes praised.[19]

Furthermore, it is not clear that one's taking an act to be supererogatory is *sufficient* for her to praise that act. Compare, on the one hand, the attitude you may have toward an act of heroism as detailed in some work of history, and on the other, your attitudes towards agents whom you admire, to whom you are grateful, and of whom you are proud.[20] Although the historical figure may in your judgment *be* praiseworthy, it would be odd if praise amounted to judging someone to have performed a supererogatory action. Just as one might say, 'I know that what he did was blameworthy, but still I can't bring myself to blame him', so too, it seems plain to me, one might unproblematically say something like, 'I know that what he did was supererogatory, but I couldn't care less', or even: 'damn him, for going over and above duty—doing one's duty is repulsive enough!' We can imagine a Satan-like figure wholeheartedly making the latter claim.[21]

One might worry that I am being uncharitable to the accountability theorist. After all, her contention is that praise is derivative of holding agents to interpersonal demands, where the latter presupposes the praiser's sensitivity to the normative force of the demands in question. It is typically held that constitutive of being an accountable agent is that one hold others and *oneself* to such demands, such that one is disposed to feel guilt or remorse

---

[19] Compare with Martin (2014: 128): "We feel appreciative when people do their jobs well, and feel grateful for even minor assistance; we even sometimes feel grateful to people who do what we believe are strict moral duties, like keeping a promise, refraining from cheating, or warning us of imminent danger. Yet we are not treating the principles governing such actions as supererogatory, because we resent failures to do one's job well, be a minimally good Samaritan, or uphold moral duties."

[20] Although I do not think attitudes of praise are easily *identified* with these reactive attitudes, they nonetheless provide a window into paradigmatic instances of praise.

[21] Here I draw on Arpaly and Schroeder (2014: 160; 161); cf. Sher (2006: 77).

when these demands are flouted.[22] Now, the Satan-like figure is the antithesis of the agent

who hold himself and others to normative demands in this sense. For this reason, the

accountability theorist may insist that the Satan-like figure does not actually recognize

another's having exceeded a normative demand. Put more generally, the objection is that

by attributing to the accountability theorist a "supererogation strategy" I am overlooking

the importance of moral psychology for the accountability theorist. That is, while it is true

that normative demands are rooted in moral requirements, on the standard way of

understanding requiredness as a feature of an act-type, an agent's acting impermissibly (i.e.

failing to do what is required) does not constitute a violation of the accountability theorist's

kind of normative demand *unless* the action reflects the agents poor *regard* or indifference

of the person he wrongs. Now, if the quality of being supererogatory— like that of being

permitted, required, or forbidden—is feature of an act-type, then the accountability theorist

does *not* think praise is properly responsive to supererogation.[23] The supererogation strategy

will fail to hit its mark if it ignores agents' quality of will.

---

[22] See especially Russell (2004), McKenna (2012: 80-8), and section 1.2 of this paper.

[23] Traditionally, a supererogatory act is thought simply to be an act that is morally good (like a duty-fulfilling act*),* but unlike a duty-fulfilling act, its performance is not required (Urmson 1958). To see that supererogation so construed is intuitively insufficient for praise, suppose that Bill has a duty to repay his neighbor $50. Suppose that, when it comes to repaying his neighbor, Bill hands him 50 dollars *and* one cent. In returning more than he owes, Bill intuitively does more than required by duty. He was not *required* to give his neighbor that extra penny, but insofar as his repaying his neighbor involves the fulfillment of a duty (and then some), which is a morally good thing, returning more than he owes seems also to be a morally good thing. In short, on a minimal view of supererogation— even on one that requires comparative cost (Benn 2017)– Bill acts supererogatorily. Yet, I submit, he is not praiseworthy. While the above considerations might seem to speak in favor of a view on which supererogatory acts involve, sacrifice (or large risks) in addition to the abovementioned elements, sacrifice-involving views of supererogation do little to establish that supererogatory acts are worthy of praise. We can imagine that when Bill gives his neighbor the $50.01 he does so knowing that he will thereby be $.01 short for his (already past-due) car payment, leading to its repossession. Many sacrifices and risks are

These are important qualifications to make. "Acting in excess of normative demand" must be understand such that the relevant demands target the agent's quality of will or regard. The core of my criticism, however (all but the point about the Satan-like figure) remains unaffected. This can be illustrated by turning to a recent and subtle theory of accountability, namely David Shoemaker's (2013; 2015). On Shoemaker's view, positive and negative accountability-responses are sensitive to a range of qualities that fall under the banner of "regard", understood in terms of the weight given to an agent's normative perspective. Accountability praise, on his view, is responsive to "*heights* [contrasted with *slights*], and these involve taking another's perspective *more* seriously than he or she may demand or expect" (Shoemaker 2015: 113). Keeping in mind the qualifications of the preceding paragraph, we can understand the idea of "heighting" another as treating them supererogatorily, so long as the "treatment" includes non-actions like affective responses. Shoemaker's (2013) more extended discussion of gratitude illustrates that his view of praise, like Darwall's, employs the supererogation strategy:

> Gratitude is not fitting for the mere meeting of demands or obligations (unless perhaps, the difficulty of doing so is significant). Instead it operates in a world of expectations, namely when one's expectations are treated like demands, or when one's mere hopes or wishes are treated like things one actually has a right to expect. The heighter takes one's normative perspective and interests quite seriously, and as with slighting, the heighter may height one either by (a) judging that the reasons provided by one's expectations or hopes count more *heavily* than even the heighted agent thinks they do, or (b) emotionally identifying with the heighted to an unexpected (but welcome) degree. (Shoemaker 2013: 117)

---

reckless or dumb, even when they are token-identical with supererogatory acts, at least as construed above. It is worth noting that on some views, praiseworthiness or 'merit-worthiness' is built into the definition of a supererogatory act (e.g. Mellema 1991: 37; Horgan and Timmons 2010: 31). To preserve the idea that supererogatory acts are praiseworthy, some philosophers hold that an act's being supererogatory requires that it was motivated by benevolent attitudes (Heyd 1982; Zimmerman 1996).

We can see Shoemaker grappling here with the task of understanding praise in a way that give demands center stage. The parenthetical caveat about difficulty, and the nod to hope suggests he is being pulled in a non-deontic direction, but he ultimately provides a picture of praise expressive of the demand prejudice. To see this, consider Shoemaker's example of the first way of "heighting" an agent (i.e. of treating an agent in a way that merits gratitude): "[s]uppose after taking up your perspective I come to fully appreciate just how much you would love to meet Tom Waits, so at great expense I arrange for it to happen. I have judged that the reasons in favor of granting your (mere) wish count for more than even you think they do" (Shoemaker 2013: 117). Shoemaker here suggests that gratitude is to be understood in terms of the relation of incongruity between the beneficiary's (A's) hopes and the benefactor's (B's) *heightened* judgment of them. We can put this incongruity in terms of a disparity between the normative weight (or reason-giving force) that A and B attribute to a hope of A's; the latter is treated by B like something A has a right to expect. One problem with this proposal is that it implies that praise presupposes error in judgment. For, if B judges A's hope to have more normative weight than does A, either B or A err. On the assumption that A is correct in judging that she cannot normatively expect others to arrange for her to meet Tom Waits, B's heighting A would require the falsity of B's judgment, specifically his judging that A's reason to see Waits is weightier than it really is. But if B is correct, matters are still worse, for this would suggest that one can height another who gives scant normative weight to their hopes, so long as one treats them as they have a right to be treated. Intuitively, however, B cannot height A simply by refraining from wronging A, notwithstanding the unfortunately low view A may have of the normative force of their hopes. This problem is an artefact of attempting to understand

interpersonal praise in terms of the same normative concepts used to elucidate interpersonal blame, namely those of normative demand or normative expectation. While it would be neat if blame was responsive to the flouting of moral demands, and praise to the exceeding of moral demands, the reality is more complicated.

I turn now to the most instructive objection to the supererogation strategy, in any of its above guises. This objection strikes at the core of the demand prejudice. Even if we grant that all and only supererogatory actions (properly construed) are praiseworthy, and that taking someone to have performed a supererogatory action either conceptually or psychologically entails praising them, the concepts of normative expectation and demand nonetheless fail to elucidate our responses of praise. To see why, note that demand plays a dual-role for blame in the accountability sense:

a) Blame represents an agent to have violated a normative demand.
b) *In blaming* an agent, one holds him to a demand/makes a demand of him.

What b) adds to a) is that concept of "demand" figures in accountability-responsibility *not only* in virtue of the fact that demand-involving actions are the *objects* of our responsibility responses; demands also typify the nature of the "moral address" directed toward the blamee. That is, to blame in the accountability sense is not only to represent an agent's action as a violation of demand (a detached judgment about a norm-violation could do *this*) but also, and arguably more importantly by the accountability theorist's lights, to make a demand of the wrongdoer (e.g. to apologize; to justify his action, to offer an excuse, etc.). For this reason, *even if* the objects of praise were constitutively demand-involving (as supererogatory responses are), we return to the point that *in praising* an agent one does not

hold her to a demand. What, however, explains (or better yet, justifies) the fact that praise does not make a demand of the praisee? I offer a response to this question in the next section by addressing a related question, namely: what *does* one do (if not make a demand) in praising another? The dual-role of demand for blame has no analogue in praise. This, together with the foregoing challenges to the supererogation strategy, provides us with strong reason to look for alternatives to the demand prejudice.

At this stage of the dialectic a certain strategy might present itself. Those who aim to understand moral responsibility in terms of interpersonal responses like the reactive attitudes might embrace the boundaries of the concept of normative demand and proceed to *restrict* the class of reactive attitudes. Jay Wallace (1994) adopts this approach. Wallace agrees that the concepts of normative demand and normative expectation are ill-suited for elucidating praise, and argues that Strawson's original list of reactive attitudes should be substantially truncated, to the following: resentment, indignation, and guilt. Wallace takes these three emotions to be central to moral responsibility in virtue of their being ways of "holding accountable", where this includes representing their target to have violated a normative demand. Wallace is not guilty of the demand prejudice: he does not assume praise is to be understood in terms of normative demand. But this achievement comes at the cost of, in effect, treating moral responsibility as exclusively about blameworthiness.

While Strawson's original list of reactive attitudes may have been too broad, we should pause to ask whether something has gone wrong if our account of moral responsibility excludes our responses toward morally praiseworthy agents. (Saints and heroes, after all, seem like the kinds of agents about which a theory of moral responsibility

should have something substantive to say.) This is especially so if upon reflection our blame and praise responses reveal a unity deeper than that found by Wallace in the concept of normative demand. In any case, if we seek to understand the nature of normative responses to a full range of responsible agents, it is incumbent upon us to enrich our vocabulary of praise.[24]

3. Estimability and What We Hope of Each Other

> *Estimable*, adj.
> mid-15c., "capable of being estimated," from Old French *estimable* and directly from Latin *aestimabilis* "valuable, estimable," from aestimare (see esteem (v.)).
> Meaning "worthy of esteem" in English from the 1690s.[25]

The aim of this section is to fill the gap highlighted by the previous section. I propose an account of responsibility in what I call the "estimability" sense. The kernel of truth in the supererogation strategy is the idea that praise is responsive to something 'above' the satisfaction of a normative requirement. Where that strategy goes wrong, however, is in construing this 'above' as belonging to the same normative register as that of requirements, namely to the deontic face of normativity. Progress can be made if we understand interpersonal praise as belonging instead to the evaluative face of normativity. Accordingly, I outline in this section the nexus of evaluative concepts proper to

---

[24] As Paul Russell (2013: 190) writes, "[a] one-sided view that is exclusively concerned with negative reactive attitudes, focusing entirely on their connection with blame and retribution, offers us an impoverished and unbalanced interpretation of responsibility and fails to properly accommodate the constructive role of reactive attitudes in endorsing and supporting morally worthy or admirable actions and traits." For another forceful challenge to Wallace's narrowing strategy, see Macnamara (2011, esp. 98-101).
[25] Harper 2017: http://www.etymonline.com/index.php?term=estimable

estimability-responsibility: normative hope; affirmation; 'holding in esteem'. In addition to providing a novel framework for understanding praise, estimability makes sense of aspects of blame that are left unexplained on the accountability model. This sections concludes with a discussion of estimability's import for understanding non-deontic blame.

Normative Hope

Return to the person who keeps her promise in the face of strong temptation to break it. Call her Alice. Had Alice broken her promise, she would have flouted a normative expectation. In that case, it would be reasonable to hold her to account for wronging you, by blaming her and thereby demanding that she repair the situation (either by apologizing; offering an excuse, etc.). But your orientation toward Alice is not exhausted by the disposition to hold her to account. After all, the gratitude felt toward Alice for keeping her promise, or *her* pride in having kept it, cannot be explained by the disposition to hold accountable. It is not as if we are grateful to everyone who fulfils a duty they have to us. I am not ordinarily grateful to the person who, instead of pushing me off my seat, remains standing on the congested subway car. Our practical orientation toward Alice, then, involves more than just the disposition to hold accountable.

In addition to holding Alice to normative requirements, we view Alice as capable of recognizing the normative force of interpersonal *ideals*. My *holding* her to such ideals, I propose, is a matter of my occupying a *hopeful* stance toward her (and toward other similar agents. There's of course nothing special about *Alice*; she's a placeholder "estimable agent", whose relevant characteristics are outlined in this section). While holding accountable is a stance of interpersonal *constraint*— wherein we stand ready to

exercise normative authority over others—the "holding" of holding estimable consists in the stance of *holding* agents *up* to interpersonal ideals. The idea that we inhabit a stance toward agents of "normative hope", alongside that of normative expectation, has recently been proposed by Adrienne Martin (2014: 130), according to whom normative hope is "like normative expectation, *sui generis* – it is its own irreducible way of relating to someone interpersonally. It is a way of treating a principle as worth aspiring to, without insisting on compliance". Since talk of "principles" arguably keeps us within the deontic domain,[26] I take the positive analogue of normative requirement to be normative *ideal*, where an ideal is, whatever else it might be, a valuable and difficult-to-achieve end that is the proper object of aspiration (see Brownlee 2010: 247). Of the ideals we aspire toward, one category concerns our treatment of, and attitudes toward, one another. Call these "interpersonal ideals". To anticipate, the content of these ideals will be constituted by the various relationships obtaining between estimable agents (e.g. friends, team-members, spouses, siblings, colleagues, strangers etc.) For an agent A to normatively hope that another agent B ϕs includes A's viewing B's ϕ-ing as fulfilling an interpersonal ideal, where the content of that ideal is—at least if A's hope is fitting— constituted (or at least, constrained and informed) by the kind of relationship that A and B find themselves in.

---

[26] This is because normative principles are typically understood on the model of *rules*, and since the rules in question cannot be binding in the way that standard normative requirements are (binding at all times and in all circumstances), it is natural to suppose that by "principles of good behavior" Martin means something like *imperfect* duties. The problem, however, is that it is unclear why the fulfilment of an imperfect duty should generate praise. It cannot be its status as a *duty*. Neither, though, does the discretionary aspect of such duties easily explain why they merit praise. That is, if I am obligated to x, the fact that circumstances of my x-ing are 'up to me' hardly elucidates why x-ing should merit praise. The explanation of the latter seems to require reference to aretaic or axiological considerations.

"Normative hope" in the first place refers to an interpersonal "stance", or disposition that structures a range of our agential responses. For this reason, normative hope is not to be identified with psychological episodes of hoping. While an agent's stance of normative hope *might* manifest itself in prospective episodes of hoping, it can manifest itself in retrospective attitudes. To illustrate, suppose that Terry's wallet falls out of his pocket as he is walking down the street. You are walking behind him and having noticed what has happened, you return the wallet. When you do so, Terry is grateful to you and says *thank you so much*! Now, Terry never had a hope with the content 'that you recover my lost wallet'. This is because, prior to your returning the wallet, Terry did not have the beliefs or desires requisite for intelligibly having such an attitude. Terry did not, for instance, believe that the wallet was missing, and so, did not desire that it be recovered. (Beliefs and desires of this sort are widely taken to be necessary, even if insufficient, components of hope.[27]) Still, in light of his standing hope (part of the complex disposition or stance of normative) that people *be good* in ways that presumptively include the type of action you perform, Terry is grateful to you. Terry's gratitude is intelligible in light of a commitment to an interpersonal ideal that can *also* be manifested in the prospective hope of e.g. the benevolent observer who watches Terry and other property-dropping pedestrians from miles away with his powerful telescope. So, although the stance of normative hope is

---

[27] On the "standard" picture of hope, hope consists in i) a desire for some state of affairs coupled with ii) a belief that the satisfaction of that desire is possible (Wheatley 1958; Downie 1963; Day 1969). It is not implausible that one must *also* engage in episodes of "mental imaging," where this requires the "devotion of mental energy to what it would be like if [the] projected state of the world were to materialize." It indeed seems essential to the hope of the aspiring gold-medal winner, say, that she picture herself on the podium, or celebrating with her supporters, or in any number of other joy-suffused scenarios (Bovens 1999: 674).

not reducible to attitudes like the benevolent observer's, it is often useful to speak of the individual normative hopes—i.e. ideal-involving attitudes of estimable agents— that either do, or would, rationalize praise-manifesting responsibility responses.

Notice that this use of "hope" mirrors the accountability theorist's use of normative "expectation". One *might* expect X in the sense of entertaining the attainment of some believed to be likely state of affairs, but normative expectation plays a structuring role in the moral psychologies of accountable agents. My resentment of the person who hurriedly shoves me upon exiting the train is explained by reference to my "normative expectation" that people respect one another's bodily autonomy (or something along these lines), but I need not have had a mental episode of expecting that this person not shove me. Rather, I occupy a stance of normative expectation, which is comprised of a range of individual normative expectation*s*, i.e., internalized interpersonal requirements/ interpersonal requirements to which one is committed.

Normative hopes, then, are internalized interpersonal ideals that figure in our practical stance toward estimable agents. The ideals in question are those that govern the relationships agents participate it. What kinds of relationships? A natural place to start would be the "moral relationship" that obtains between members of what Strawsonians call the moral community. Scanlon (2008; 2013) offers an account of blame that is based upon the reasons generated by one's involvement in relationships of this sort, where a relationship is just "a set of intentions and expectations about our actions and attitudes toward one another that are justified by certain facts about us" (Scanlon 2013: 86). To blame some person P, on Scanlon's view, is to *modify* one's intentions regarding, and

expectations of, P in light of one's judgment that P is blameworthy.[28] Regardless of whether Scanlon provides a satisfactory picture of the nature of blame, there is intuitive promise to the idea that our responses of praise and blame should be understood in terms of the norms governing the kinds of relationships that unite agents.[29] A virtue of this approach is that it provides a natural way of understanding why different forms and degrees of blame and praise are appropriate toward persons with whom we stand in different relationships, e.g. that of friend, colleague, spouse, teammate, stranger, advisor, etc. The norms governing each kind of relationship will be differ, and so, the type of blame appropriate in each will understandably differ. I propose, however, that in addition to expectations and intentions— the attitudes Scanlon identifies as relationship-defining— we include hopes. Among the norms that govern our relationships are plausibly ideals that give us reasons to hope that those to whom we are variously related *live up* to our relationship-defining ideals. This is a revision Scanlon may be happy to accept, for he holds that our relationships are assessable relative to their respective normative ideals (rather than merely to a set of requirements).[30]

---

[28] "To *blame* a person is to judge that person to be blameworthy and, as a consequence, to modify one's understanding of one's relationship with that person (that is, to alter or withhold intentions and expectations that that relationship would normally involve) in the particular ways that that judgment of blameworthiness makes appropriate, given one's relation with the person and the significance for one of what the person has done" (Scanlon 2013: 89).

[29] I adopt a weaker thesis about the relation between relationships and norms than does Scanlon, who takes the relevant norms to be *constitutive* of the relationships in question.

[30] "We have not only intentions and expectations regarding our interactions with friends and associates but also intentions and expectations that define a relationship with other people in general. We have, for example, views about the intentions and the care one should take not to injure strangers and the duties one has to aid them should we be in a position to do so. *The normative ideal of the moral relationship* specifies what these intentions and expectations ought to be. *Our actual attitudes may fall short of this ideal*, but we have them nonetheless." (Scanlon 2013: 88, italics added)

Scanlon's focus on expectations and intentions is likely explained in part by the fact that his discussion of relationships is introduced in the service of fashioning an account of blame, where normative expectations are more plausibly central. A description of one's relationships to strangers, friends, family-members, and colleagues, to be sure, will require reference to the intentions and expectations she has towards these persons, but so too will it require reference to the hopes she has of them. By allowing that hopes are among the kinds of attitudes governing our relationships we make available the following, quasi-Scanlonian, description of an important manifestation of praise, namely *relationship-enhancement*:

> Praise of a person ordinarily involves judging a person to be praiseworthy and being disposed to modify one's understanding of one's relationship with that person (that is to alter or extend intentions, expectations *and hopes* that that relationship would normally involve) in the particular ways that that judgment of praiseworthiness makes appropriate, given one's relations with the person and the significance for one of what the person has done.[31]

While we should avoid *identifying* praise with relationship-enhancement—for one, praise of the deceased would be rendered mysterious— it is no accident that we tend to grow warmer toward those whom we have occasion to be proud of, admire, and be grateful to. This means, among other things, that we hope that they *do well*, and that things go well for them (and may be let down, disappointed, or otherwise saddened, when they do not).

---

[31] This is of course a modified version of Scanlon's characterization of blame (see the previous footnote). I say that it is quasi-Scanlonian, however, because a) it is offered not as a definition of praise and b) Scanlon claims that "praise is a purely evaluative notion," by which he seems to mean that it consists in the sort of assessment covered by attributability-responsibility (Scanlon 2013: 86; 95). Slightly confusing is that Scanlon takes gratitude to be a positive reactive attitude (Scanlon 2013: 96-99). I'm not sure what he takes this to be if not a vehicle of praise.

This suite of hopeful attitudes is not limited to interchanges between participants of intimate relations. After all, it has application in the moral relationship obtaining between all agents possessed of the capacity to recognize interpersonal ideals. Consider an intermediary case, one obtaining neither between friends nor complete strangers. Suppose Mel and Pam are co-workers. Mel overhears Pam, recently hired, expressing frustration about their employer's unwillingness to give her a certain day off so that Pam can attend a performance by her favorite musician. Although it would be slightly inconvenient for Mel to take the shift, she is free that day and after briefly deliberating and empathizing, she offers to swap shifts with Pam. Pam is grateful; her friend admires Mel for her act of kindness; Mel feels slightly proud of herself. Each of these attitudes is naturally explained by a common normative hope, one to the effect that people help one another when the cost of doing so is little to nothing. Mel's action is presumably an instance of this estimable action-type. Even if none of the parties in fact prospectively hoped that Mel would φ, φ-ing is, by each of their lights, 'something to be hoped for', i.e., the kind of thing one would hope people would do for one another. Granted, what Mel did may be supererogatory, but this fact provides little explanation of why she merits praise. If Pam had taken out a debt-incurring loan to hire Mel's favorite band for an exclusive concert, Pam would have acted supererogatorily, but the excessiveness of the act at least detracts from its admirableness; such actions are not obviously commended by the ideals governing their relationship.[32] While we might be able to think of a kind of relationship in which such a grand gesture would be, that's just not the sort of thing *co-workers* (qua coworkers) do for one another.

---

[32] And again, the feature of supererogatoriness is missing from cases in which persons are praised for doing what can be normatively expected of them.

In addition to the normative hopes that structure our relations and responses to all persons insofar as they are agents capable of responding to interpersonal ideals, we do have hopes that define our personal relationships. "Personal hopes" are not only *directed towards*, but also *invested in* particular persons. To see what this contrast comes to, consider the attitudes and tendencies a piano teacher might have toward her pupil. On the basis of considerations like their history before the piano, the teacher's estimation of her pupil's potential, the pupil's passion and willingness to master her instrument, etc., the teacher might invest hope in the student to master her instrument. While the teacher can recognize the value of *other* students becoming excellent pianists, the satisfaction of her normative hope requires that *she teach this student*— her pupil— to recognize the musical value of which she is, hopefully, capable. Personal hope provides a way of understanding a form of second-personal praise that is overlooked in the literature on the reactive attitudes, namely second-personal *pride*. This is the attitude the piano teacher expresses in saying "I'm proud of you".[33] In saying this, the teacher is affirming the value of what the student has done (or become), something that a stranger may be in a position to do, but only impersonally, i.e. out of an attitude of admiration. But this brings us to a second facet of responsibility in the estimability sense.

Affirmation and Holding in Esteem

An upshot of my proposal is that our moral community is not only one of accountable agents, but of estimable agents too. We hold others estimable by viewing them from the

---

[33] The widespread assumption that pride is essentially a reflexive emotion is also rejected by Helm (2017: 49).

stance of normative hope, i.e. as beings capable of aspiring toward ideals reflected in our interpersonal normative hopes. When an agent satisfies a particular normative hope, his interests are apt to be positively affected by other estimable agents (especially the person he *heights*). In this case, he might actually be *held in esteem*, by e.g. being the target of praise-manifesting attitudes (like gratitude, admiration, and pride) and actions, as well as through modifications in the praiser's good will and trust. This brings us to the communicative character of estimability.

Recall that reactive attitudes are often taken to be incipient forms of moral address. I stick here to *communicated* reactive attitudes, as these will be our clearest guides to the kind of address that the attitudes themselves might 'aim toward' or, incipiently *be* (Watson 1987: 264). The address in question is typically understood in terms of moral *demand*. I have argued that although demand may characterize the form of address proper to negative reactive attitudes, it has no place in describing the positive reactive attitudes. What kind of address, then, is involved in praise? My proposal: *affirmation*. Consider the kind of response characteristic of the grateful agent, namely expression of *thanks*. To thank a benefactor is to address her for what she has done, but not in a way that even implicitly makes a demand of her. In thanking another I do more than *assert* or imply that she has done something good for me. (The enemy of all that is good can acknowledge that what one did was, say, morally good.) In addition to its assertoric function, praise has an *expressive* function, which is explained by its being expressive of the *satisfaction* of a desiderative attitude, namely a normative hope (understood as a feature of the praiser's

psychology).[34] That is, communicated praise *affirms*. While we do not typically refer to 'satisfaction' as a mental state alongside desire and belief, we know what it is for a desire to be subjectively satisfied. Typically it feels good. When the satisfied desire reflects a normative hope, it will necessarily be satisfied by an agent's acting well. Naturally, then, the joy taken in the desire's satisfaction will be directed toward the agent. In thanking a benefactor, for instance, one joyfully affirms the value of what she has done. In so doing, one not only asserts that something desirable has been done; furthermore, one overtly *upholds* it as (non-instrumentally) valuable and, since one is the beneficiary, *welcomes* what was done.[35] Praise, in this sense, says "yes". It affirms.

When one praises another as an observer (rather than as a beneficiary), one may express admiration for what another has done. Here too one affirms as valuable what an agent has done. As Macnamara (2011: 89) writes, in discussing expression of positive reactive attitudes, "I express my gratitude in a heart-felt thank you or my approval in words of recognition: ''That was a truly lovely thing to do.''" When they are uttered by an agent who admires (and so, values) what was done, the words "that was a truly lovely thing to do" do not only (or primarily) ascribe the property of "loveliness" to an act; they express the praiser's being *for,* in the sense of upholding as (non-instrumentally) valuable, what was done. The proposal at hand helps explain the failure of attempts to deflect praise with

---

[34] This outline of affirmation stands in contrast to that recently proposed by Wallace (2014), according to which affirmation is an intention-like attitude in that, to affirm x is to intend to bring x about *if one could* (Wallace 2014: 62). On my view affirmation, like hope, is *desire-like*, and the same kinds of desires can figure in both. While hope represents the object of a desire as possible/attainable, affirmation represents the object of one's (past) desire to be (or have been) actual/attained.

[35] It is perhaps possible for one to feel and communicate gratitude against one's best judgement, but I restrict myself here to cases in which one's normative hopes correspond with her judgments about what is valuable.

responses like "I was only doing my duty". (We can imagine the lifeguard saying this in response to the praise of a grateful parent.) By persisting in the face of such a reply, praise seems to say, "that may be so, but I am here to *affirm the value* of what you have (even if dutifully) done."[36]

Overt expressions of affirmation, to be sure, do not exhaust the ways in which praise-manifesting attitudes can address their targets. *Warm behavior* refers to a category of less agentially robust expressions of praise, e.g. gestures of friendliness, increased trust and good will, greater sympathy, greater attention to the welfare of the person to whom one is grateful, subtle emulation of the object of admiration, etc. Warm behavior is the more covert cousin of overt affirmation, for it includes responses toward the praisee that might escape the notice of the person praising. Although they are not ways of explicitly addressing the person praised for what she has done, in order to count as expressions of praise, warm behavior must be based in belief (or some similar state) about the praiseworthiness of the person praised.

Estimable Agency & Non-Deontic Blame

Responsibility in the estimability sense has the further virtue of shedding light on a poorly understood type of blame, namely blame for treatment that is *bad* but does not deprive others of what they're *owed*. To give others what they're owed—i.e., to respect the claims

---

[36] A further virtue of the proposal is that it provides a way of understanding *prospective* forms of address to estimable agents. To return to our piano teacher and student, it would be natural for the teacher to express the hope she invests in her student by offering words of encouragement. Although words of encouragement are not ways of praising, they affirm the student in a forward-looking way by expressing emotional support for her pursuit of a jointly valued end.

that others have as moral agents—is what accountability, as it were, demands. But there are ways of treating others well and poorly that do not fall under the scope of what they're owed, and of what can be demanded of us. Consider Scanlon's characteristic claim that, "we do not owe it to people generally to trust them, be willing to cooperate with them" (Scanlon 2013: 91). Darwall similarly claims that "[s]omeone who fails to esteem your estimable qualities may not give you the response you deserve, but esteem is nothing you or anyone else can expect or demand" (Darwall 2006: 120). On the face of it, an agent might never hold others or himself to interpersonal ideals, or act estimably, yet satisfy the requirements of minimal moral decency. I have in mind an agent who is *capable* of holding agents to interpersonal ideals, but who refuses to adopt a moral stance beyond that of recognizing and responding to normative demands. He gives other what they're owed, no more and no less. The more we think about what would be missing in the life of such an agent, the more we can see that he would strike us as normatively defective. This may be because there is a notion of "common decency" that falls within the domain of estimability. Consider Cheshire Calhoun's description of Scrooge, a character who lacks "common decency": "That Scrooge has no truck with simple favor, such as telling others the time or giving them directions, signals his lack of common decency" (Calhoun 2004: 128). Though he is the object of the reader's contempt (and perhaps pity)—and Scrooge himself comes to be ashamed of his stance toward others—Scrooge's lack of common decency is intuitively not explained by the *accountability theorist's* standard of 'minimal moral decency', for Scrooge doesn't obviously flout normative requirements. As Calhoun puts it:

> for all that Scrooge is an "odious, stingy, hard, unfeeling man," there is no one whom he clearly *wrongs*…Dickens gives us no reason to think that Scrooge fails to render what is due. On the contrary, Scrooge is obsessed with debts. He wants nothing more from others than exactly what they owe him. In return, he will give

others exactly what he owes and not a bit more. His business and moral ledgers carefully track debts payable to and by him, making no allowances for giving or receiving that exceeds the obligatory. And this is the source of his failure of common decency. (Calhoun 2008: 129)

Suppose Scrooge, having read Darwall and Scanlon, responds as follows to all requests for favors, beneficence, mercy, forgiveness, volunteering, and the like: 'I don't owe it to you to do that!' What should an adequate theory of moral responsibility say about him? On the one hand, he is right: so long as he acts in such a way that does not *wrong* another, Scrooge gives him what he owes. On the other hand, however, we think Scrooge behaves contemptibly. The thought underlying this puzzling assessment is captured in Robert Adams's claim that "we ought in general be treated better than we deserve" (Adams 1985: 24). If this means 'we are owed more than what we are owed', or that 'we have a right to treatment to which we have no right' the claim is incoherent. Incoherence of this sort is difficult to avoid if we remain within the deontic realm in parsing, 'we ought in general be treated better than we deserve.' My proposal is to read the above "ought" as one of estimability. Doing so provides an explanation of why we may be justified in blaming those who have not obviously violated any moral requirement.

While many responsibility theorists hold that blame is a response to the violation of normative requirements or duties, others have proposed that we can appropriately be blamed not only for wrong action of this sort, but bad actions that aren't wrong. McKenna, for instance, denies the "restrictive" view of blame, on which blame is properly responsive only to violations of obligations: "I fail to see precisely why an agent could not be blameworthy for acts that are criticizable in exclusively axiological or instead exclusively

aretaic rather than deontic terms." (McKenna 2012: 182-3).[37] Putting aside the aretaic realm (as that arguably puts us in the territory of attributability), let us consider what may be involved in blaming a person for bad but not wrong action, or "suberogatory" action (Driver 1992). I borrow the example from Driver (1992), via McKenna (2012: 184): "Bob is suffering from kidney failure, and while Roger is not morally obligated to give Bob a kidney, his failure to do so would reveal a failure to live up to a moral ideal. About this case, Driver writes:

> Giving up one's kidney to save one's brother is certainly acting according to a moral ideal—fraternal self-sacrifice. Yet, it conflicts with the right one has to keep one's own kidney. Note that now failing to act on a moral ideal does seem blameworthy, even though permissible. ([Driver 1992:] 287–288)"

It is not my aim here to defend the class of suberogatory actions. Still, there is considerable intuitive appeal to the idea that we can be blameworthy behavior that does not constitute *wrongful* action, or violations of obligation. What is peculiar about this class of blameworthy actions is that a) the agent in question seems to be blameworthy in more than just the attributability sense (suppose her bad action is out of character), and b) the agent is not accountable for such action, at least not according to the standard way of understanding accountability. Now, in rejecting the restrictive view McKenna proposes that an agent *can* be accountable for bad action that does not amount to the violation of a normative requirement. The difficulty with this idea is that it is not clear what could unify accountability as a type of responsibility. McKenna treats "normative demand" as the

---

[37] See also Macnamara (2011: 88): "what reason is there to think that reproving another for her wrongdoing is an instance of holding responsible, but criticizing another's bad conduct isn't? Responsible agents are, after all, just as responsible for their bad actions as they are for those that are wrong. It thus seems odd that we hold others responsible for the latter but not the former."

concept focal to accountability, but if this cannot mean "normative requirement", accountability blame begins to look mysterious.

Estimability, I propose, has the resources available to explain the nature and appropriateness of axiological blame (that is, blame for what's bad but not wrong). I do not pretend it is the only explanation, but it has the virtue of shedding light on how this sort of blame is and is not like the standard accountability theorist's (deontic) blame. When I act badly toward someone, although I do not 'wrong' them' or violate a demand of justice, I may *slight* them, where this is a matter of failing to treat them with esteem. I do something I ought not have done, where this is an 'ought' not of moral obligation, but of an interpersonal ideal. If 'oughts' of the latter sort are sometimes experienced as having binding normative force, this may be because estimable agents expect more of one another than what justice strictly demands. Responsibility in the estimability sense thus provides a non-paradoxical way of understanding Adams's (1985: 24) claim that "we ought in general be treated better than we deserve". There is a sense in which morality requires that we not only *do right* by one another, but that we be good to one another. This thought, anyway, is reflected in our responsibility practices. To say this is not to espouse universal or unconditional benevolence, which would be to undercut the importance of interpersonal desert. To questions of estimability and desert I now turn.

4. Deserved Moral Address:
Assessing the Prospects for a Unified View

Having rejected the accountability theorist's proposal of understanding interpersonal blame and praise in terms of normative demand, the following question naturally arises:

what, if anything, unifies interpersonal praise and blame? *Is there* an underlying unity to accountability and estimability? Progress on the 'unity question' can be made by focusing on a pair of features common to estimability praise and accountability blame. First, both can be understood as forms of "moral address", which at a minimum means that they have the function of seeking uptake in their targets. This address has the form of "demand" for blame, and "affirmation" for praise. These forms of address share an important characteristic, bringing us to a second feature common to accountability blame and estimability praise, namely their *interest affecting* character. That is, both are evaluative forms of address that by their nature characteristically affect the interests of their targets. First, blame characteristically adversely affects one's interests— or *harms* the blamee (Feinberg 1986; McKenna 2012: 134-141). It is in virtue of this connection between blame and harm that Watson (2004 [1996]: 74) and Wallace (1994: 275; 280) claim that blame involves liability to sanctions.[38] The idea here is not that A's blame of B *necessarily* harms B. (The key word is "liability".) After all, A might blame B privately; B might fall outside A's sphere of causal influence; B might fail to understand A's expression of blame as a moral demand, etc. Nevertheless, there is for the accountability theorist more than a *merely* contingent connection between being blamed and being harmed. The blame in question is part of a practice of holding responsible that is inherently interpersonal. In *paradigmatic* cases of blame, the blamer communicates her opposition to the blameworthy action by

---

[38] "Holding accountable…involves the liability to sanctions. To be entitled to make demands, then, is to be entitled to impose conditions of liability," for, "blaming responses (at least potentially) affect the interests of their objects adversely" (Watson 2004 [1996]: 275; 280); "True moral blame…is a form of deep assessment, reflecting an attitude towards the agent that finds its natural expression in sanctioning behavior (avoidance, denunciation, reproach, censure, and the like)." Wallace (1994: 74).

overtly making a demand of the blamee (to offer justification or excuse; to acknowledge wrongdoing; to apologize; to rectify the situation, etc.). Expressed blame of this sort is what the wrongdoer will, absent countervailing reasons, justifiably (or deservedly) receive in virtue of being blameworthy.

In saying that blame is *paradigmatically* expressed blame I do not mean that communicated blame is, in all respects, explanatorily prior to non-paradigmatic instances of blame.[39] The representational content and valence of the reactive attitudes of blame, importantly, are not obviously explained by their overt manifestations. I labor this point partly because I do not wish to deny that private blame is *true blame*, itself governable by norms of appropriateness. Overt blame is of special practical significance, however, owing to its role in facilitating moral repair, growth, and the wrongdoer's reintegration into the moral community. Reintegration of this sort is often morally desirable, but this does not mean that blame that precludes reintegration of this sort (either by remaining private or by foreswearing the possibility of forgiveness) is for that reason not fully blame. It is arguably an unfortunate feature of some cases of blame, e.g. blame of the deceased or other unreachable agents, that since the address can only be *hypothetical*, the prospect of reconciliation remains hypothetical too. None of this challenges the idea that reactive attitudes of blame are communicative entities in virtue of their function of demanding from their target a certain uptake.[40] The link between blame and harm plausibly, then, comes to

---

[39] McKenna (2012) commits himself to this claim, the "publicity thesis". For a significant challenge to the publicity thesis see Russell (2017b).

[40] The term "communicative entity", I borrow from Macanamara (2013; 2015). On her view, reactive attitudes are like email messages in being "intrinsically directed toward a specific end—the end of eliciting uptake of their representational content in a recipient." (Macnamara 2015: 218) Unlike other messages, however, a reactive attitude is according to Macnamara "a communicative entity that takes the wrongdoer as its addressee, i.e., the

this: in virtue of our being the kinds of beings we are (particularly, those with the moral psychological profiles of accountable agents), our interests tend to be adversely affected by being targeted with the demands of blame (and absent countervailing reasons, are *apt* to be so affected). These demands are often emotionally painful, even if we recognize them to be fairly imposed.

Praise, by contrast, characteristically *benefits* the praisee. That is, being praised favorably affects the interests of the praised. How so? The awardee of a Nobel Prize is in an obvious sense benefitted, but reward of this sort—which is instituted with *the goal* of benefitting the recipient—is inessential to praise. Since "benefit" here means something that promotes one's interests, we need to clarify what interests are at issue. I focus here on "welfare interests",[41] which, following Feinberg (1986) I take to include "the absence of groundless anxieties and resentments, the capacity to engage normally in social intercourse and maintain friendships, at least minimal income and financial security, a tolerable social and physical environment, and a certain amount of freedom from interference and coercion" (Feinberg 1986: 37). Now, Feinberg characterizes welfare interests to set a *minimal* threshold of goods requisite, by themselves insufficient, for a good life (Feinberg

---

person from whom they seek uptake, where uptake amounts, at least in part, to sincerely acknowledging one's fault." (Macnamara 2015: 222)

[41] While I do not here consider the possibility that one can be harmed or benefitted by another's private/non-communicated attitudes of blame or praise, I do not mean rule it out. There is some plausibility to the idea that B might be harmed by A's private blame, e.g. because A's blaming B (even privately) reduces the beneficial *possibilities* that B might have open to him. The idea that private blame may harm its target rarely gets extensive defense, but sympathy for it is found in Nelkin and McKenna. Nelkin (2013: 124) writes, "I find it plausible that things can be bad or good for us without our being aware of them. So it may very well be that your resenting me is bad for me even if I never find out about it and am not affected by it in indirect ways such as your spending less time with me." And McKenna (2012: 132 n. 8): "It is probably true that a person could be harmed by another's negative regard for her, even if the other's regard were kept private."

1987: 37). But these welfare interests can easily be 'scaled-up' to shed light on what it is to be benefited. For instance, improvements in one's social environment—including those that make one's social environment more than "tolerable"— are naturally understood as promoting one's welfare interests, and so constitute benefits. The relationship-enhancing manifestations of praise intuitively comprise this sort of good. As noted above, we tend to grow warmer and closer to those of whom we are proud, to whom we are grateful, and whom we admire. While blame manifests itself through the alienation and exclusion of the blamee (McKenna 2012: 137-139; C. Bennett 2002: 151-2), expressed praise is welcoming, and tends toward solidarity, or the praisee's greater inclusion in the praiser's region of the normative community. Praise thus tends to have a favorable effect on the praisee's social relations. Next, one's emotional landscape might be not only free from "groundless anxieties and resentments", but also inhabited by warranted pleasurable emotions, which may have their bases in the affirmation characteristic of praise. Expressed praise involves responding positively to the praisee, and so is poised to please the praisee by inviting attitudes like pride, fellow-feeling, and an increased sense of purpose.[42]

To nearly return to the Nobel Prize recipient, even if we reject the Sidgwickian view that reward is gratitude universalized (bk. III, ch. 5, 281), there remains a natural connection between praise (including praise-manifesting emotions like gratitude and admiration) and generous actions that are reward-like. Although in the interpersonal

---

[42] There is a case to be made for the thesis that praiseworthiness ought to be understood in terms of the desert of some such response, perhaps most plausibly what we might call "agential pride" (the positive analogue of guilt). This proposal earns some support from recent efforts to understand blameworthiness in terms of deserved guilt (see Brekke-Carlsson 2016). Whether 'deserved self-blame (i.e. guilt)/self-praise (i.e. pride)' views are in competition with 'deserved address' views depends on the intelligibility of the idea of affective self-address, which is a topic for another time.

domain these actions are not always distinguishable from *gifts*, which though they may be given in recognition of praiseworthy actions or tendencies, are not to be described as instances of "rewarding" another. While I *might* be grateful to and praise the individual who responds in the desired way to my '*lost dog*—$300 reward' poster, the compensatory core of the reward remains even after this person extinguishes my gratitude to him (but not my joy that my dog is returned) by admitting that he returned the pet *solely* for remunerative reasons and would not have returned the pet had the reward been a dollar lower, despite living one street away. Estimability, then, opens up estimable agents to treatment that is reward-*like*.

Importantly, while it is plausibly among the aims of reward to benefit the recipient, I do not need to *intend to* benefit another for my praise to benefit him. Recognizing the courage required for Alex to do what he did, I might express my admiration without setting out to 'make him feel good' or to further any of his interests. And, I certainly do not generally intend to adopt a warm, pride-inducing, disposition toward those to whom I am grateful, whom I admire, and of whom I am proud.

Let's take stock. Accountability blame and estimability praise are both paradigmatically expressed through moral address (demand and affirmation, respectively), which forms of address characteristically impact their targets' interests, adversely or favorably; blame tends to harm the blamed, and praise tends to benefit the praised. Now, if some member of the moral community is to be harmed or benefited, the prospect of their being so treated raises questions regarding desert. Owing to their tendencies to harm and benefit their targets, responses of blame and praise in the accountability and estimability senses are essentially responses that are evaluable as deserved or deserved (and perhaps

(un)deserved *to a degree*, but I put aside this issue for simplicity's sake). Blame's warrant is often articulated in terms of desert; for the blame of some agent to be warranted is for that agent to, in some sense, *deserve* the harm of blame. While our attributability responses may be assessable for *fittingness*—the positive appraisal of Al as generous will be fitting just in case it correctly represents Al (see D'Arms and Jacobson 2000) —our more robust responsibility responses, those central to accountability and estimability, are assessable for their *deservingness*. For, one's being praiseworthy or blameworthy in the estimability and accountability sense opens one up to benefits or harms in virtue of giving *others* defeasible reason to treat the deserving agent in ways corresponding to the attitudes deserved (Carlsson 2017: 112; Pereboom 2017: 131; Clarke 2013; 2016).[43]

What factors determine whether an agent deserves to be praise or blamed? It will be useful here to briefly set out some terminology concerning desert. First in discussing desert I restrict myself to claims where *persons* are the prospective "deservers". Next, a "desert claim" is a proposition to the effect that some person deserves something. The thing deserved (in a broad sense of 'thing') is the "desert", though since praise and blame are our subject, I will refer more narrowly to "deserved response". Next, a desert claim requires a "desert base", that in virtue of which one is claimed to deserve the deserved response

---

[43] On this understanding of desert, a fitting attitude can be characteristically (or even inherently) painful or pain-inducing (and so, harmful on many views of harm), without being deserved (e.g. grief, embarrassment, regret, disappointment, and heartache) (Carlsson 2017: 110-113; Pereboom 2017: 131; cf. Bok 1998: 168-9). The same point holds of pleasurable (and so, beneficial) attitudes like joy, relief, love, and (some kinds of pride). These considerations speak in favor of a view on which the relevant kind of desert is not reducible to fittingness (King 2012; Nelkin 2016: 178), even if the desert-relation is a species of the fittingness-relation. The 'aestheticization' of desert can in this way be resisted (see Feinberg 1970: 82; Sher 1987: 114).

(Feinberg 1970: 58). What kind of thing *is* a desert base? It can be understood as a property of the deserver, e.g. the property of having acted (or omitted) culpably or admirably.[44]

Our foray into desert terminology can now be put to use. Recall that this section's aim is to characterize, in general terms, a type of responsibility that plausibly includes accountability and estimability as subtypes. Here is my proposal: accountability and estimability are both forms of responsibility in the *deserved moral address* sense. Responsibility in the deserved moral address sense (R-DMAS) is silent on the question of what *ultimately justifies* a desert claim. To illustrate, take the following desert claim: 'Uri deserves praise [/affirmation] for having gone out of his way to return the stranger's wallet.' That which is proposed to locally justify this desert claim, the desert base (="X," i.e. having gone out of his way to return the stranger's wallet) is apt to serve as a desert base whether one take the desert of R-DMAS to be basic (i.e. normatively primitive) or non-basic. To clarify, let's start with basic desert.

If the desert claims proper to R-DMAS (and by extension, estimability) were those of *basic* desert, the desert base would, all by itself, justify the desert claim. That is, Uri would be deserving of praise [/affirmation] *just because* of X. Furthermore, if R-DMAS is a type of responsibility in the basic-desert entailing sense, contractualist or forward-looking considerations could have no bearing on whether Uri deserves praise.[45] Responsibility in the basic-desert entailing sense might for this reason be compared with retributive justification for reward or punishment, although it is a mistake to subsume praise and blame

---

[44] Alternatively, desert bases may be construed (not as properties of deservers) but as *facts* (e.g. the fact that Jane acted admirably); little hangs on which we settle on (see Feldman and Skow 2015).

[45] Though they well might bear on whether we have reason, *all things considered,* to overtly praise Uri.

under retributive justice (as does Glover (1970) for example), at least if retribution includes

the *goal* of causing suffering or pleasure.[46] In the literature on moral responsibility and basic

desert, the terms of the debate are captured by Pereboom, according to whom:

> For an agent to be morally responsible for an action in this sense is for it to be hers
> in such a way that she would deserve to be blamed if she understood that it was
> morally wrong, and she would deserve to be praised if she understood that it was
> morally exemplary. The desert at issue here is basic in the sense that the agent
> would deserve to be blamed or praised just because she has performed the action,
> given an understanding of its moral status, and not, for example, merely by virtue
> of consequentialist or contractualist considerations. (Pereboom 2014: 2; Pereboom
> 2001; Pereboom 2007).

Now, the accountability theorist who thinks we can deserve blame in the above sense may

extend this commitment to praise in the estimability sense. This can be done without

assuming that the deserved responses and desert bases for praise are to be understood in

terms of "normative demand".

Estimability praise, however is neutral on the whether the desert base is to be

construed in the sense of basic or non-basic desert. Fortunately so, since the accountability

theorist need not think we ever deserve blame in the basic desert sense. Manuel Vargas

(2007) proposes a view of interpersonal moral responsibility on which our practices of

praise and blame are ultimately justified by their consequences. Nonetheless, he accepts

that desert bases serve as adequate *local justifiers* of desert claims about praise and blame.

That is, on Vargas's view, "particular assessments of responsibility will rarely if ever attend

to justification for the practice. Instead, it will straightaway appeal to a norm of praise or

blame. Nevertheless, what [ultimately] justifies those norms and attendant practices are its

---

[46] As Glover (1970: 72) writes, "[r]etributive justice concerns giving people what they
deserve, but they deserve praise or reward, as well as blame or punishment."

effects on getting us to attend to moral considerations." (Vargas 2007: 158)[47] For Vargas, then, Uri will deserve praise in virtue of his having gone out of his way to return the stranger's wallet praise (though the justificatory force of this desert base will be borrowed from the moral effects of the network of desert claims and deserved responses). To modify an example of Vargas's, both you (the conventionalist about baseball rules, I hope) and the divine-command theorist about baseball norms can agree that the batter is 'out' in virtue of his three strikes. This is so despite of the fact that you'll disagree about what *ultimately* justifies the norms of the game.

So far, I have restricted myself to the topic of unifying estimability with accountability. The latter, after all, is the dominant theory of interpersonal moral responsibility, and estimability earns some of its immediate appeal from the shortcomings of the accountability theorist's central concepts. Still, estimability does not presuppose accountability. The two are not a package deal. An interpersonal account of blame need not identify normative demand as its characteristic form of address. There are communicative views of blame identify blame with a form of address other than demand. Consider philosophers who understand blame as "protest" (Hieronymi 2001; Talbert 2012; Smith 2013; Pereboom 2017). Hieronymi (2001: 530) for example writes, "[r]esentment…should be understood as protest. In resentment the victim protests the trespass, affirming both its wrongfulness and the moral significance of both herself and the

---

[47] Also: "Whether an action deserves praise or blame depends on (1) whether the agent that did it is appropriately subject to norms of responsibility (young infants are presumably not, normal mature adults presumably are), and (2) what the norms of praise and blame say about actions of that sort in the relevant sort of context." (Vargas 2007: 154)

offender." Smith (2012: 39). similarly claims that "[t]o morally blame another… is to register in some significant way one's moral protest of that agent's treatment of oneself or others." To say that protest views of blame treat blame as communicative is to note that, as with demand theories, proponents of protest theories hold that blame "seeks some kind of moral reply".[48]

Like demand theorists of blame, protest theorists tend to be compatibilists. Derk Pereboom is an exception. Although he maintains that the moral responsibility presupposed by the *demand* involving blame (which he identifies with the blame of the reactive attitudes) is unwarranted (given the kinds of agents we are), there nonetheless remains a type of blame that, according to Pereboom, is immune to the threat of determinism (and, for that matter, indeterminism). Pereboom notably develops a protest-view of blame that is ultimately justified by forward-looking considerations: "Blaming is…justified by these forward-looking elements [i.e. protection, reconciliation, and moral formation] [where] the notion of moral protest is the essence of blame in the sense I'm setting out." (Pereboom 2017: 125; 128). Blame, in the form of moral protest is on Pereboom's view justified by axiological considerations stemming from its consequences, namely the goodness of the following "moral desiderata: protection of potential victims, reconciliation to relationships both personal and with the moral community more generally, and moral formation." (Pereboom 2017: 134). Still, like Vargas's, Pereboom's view of blame is not thoroughgoingly forward-looking. Both differ from the act consequentialist's theory of

---

[48] There are important differences among these views. Talbert and Smith, notably, deny that the wrongdoer need be the addressee of blame. For discussion and criticism of the idea blame's communicative function includes uptake sought in persons *other than* the wrongdoer, see Macnamara (2015: 222-228).

blame, on which a person ought to be blamed just in case, and *irrespective of their culpability or innocence*, their being blamed will result in e.g. the greatest aggregate welfare. That is, for Pereboom desert does not wholly drop out of the picture (though *basic desert* does). The desert base remains an essential backward-looking consideration in virtue of which blame is locally justified. It is for this reason that Pereboom's view is only *largely* forward-looking:

> In my account, blame as moral protest is to be justified in the largely forward-looking wat I've specified, by the aims of protection, reconciliation, and moral formation. But the immediate target of blame is typically a past action, and in this respect such blaming will have a backward-looking aspect: the badness of the past act is part of what makes the protest appropriate. (Pereboom 2017: 130)

This backward-looking aspect may be the same desert base that, for the basic desert theorist, does all the normative work.[49] A desert base of this sort may be provided by Uri's

---

[49] Though unless the basic desert theorist understands blame as protest, they are nonetheless bound to disagree about the *deserved response*. Crucially, guilt is *not* deserved according to Pereboom. It is central to Pereboom's hard incompatibilism that we are not agents who can deserve (either basically or non-basically) reactive attitudes. One might for this reason wonder whether it is right to ascribe to him a type of non-basic desert view of blame. After all, instead of guilt (which basic desert theorists [e.g. Clarke 2013; 2016] and non-basic deserts theorists [e.g. Vargas 2007] tend to identify as the appropriate response the blameworthy agent is to have in recognition of her wrongdoing, Pereboom identifies regret as the right attitude. Regret, importantly, may be fitting or unfitting, but is intuitively not the sort of attitude that need be deserved or undeserved (see fn. 38) (Pereboom 2017: 131; also, Carlsson 2017: 110-113). But, while regret isn't constitutively assessable for desert, it's less obvious that this is true of protest. That Pereboom's analogue of self-blame (namely, regret) is assessable for fittingness but not deservingness leaves open that the protest-blame of thers may be non-basically deserved, such that others have defeasible reason (which reason will presumably borrow much of its normative force from the attractive consequences Pereboom identifies) to impose/issue blame (in the form of protest), for one's morally bad action. Pereboom owes us a fuller account of what protest is, given that other protest theorists take it to be resentment.

having gone out his way to return the stranger's wallet (at least assuming Uri acted in light of an interpersonal ideal).[50]

It is open to Pereboom to supplement his axiological account of blame as moral protest with an account of praise as moral affirmation. Neither kind of blame or praise may be basically deserved (owing to the kinds of agents we are), but these responses may nonetheless be ultimately justified by their valuable effects (in the case of praise, candidates are: benefitting agents' welfare interests; enhancing, by affirming the value of relationships both personal and with the moral community more generally, and moral formation'), and locally justified by their desert bases, i.e. the properties of agents in virtue of which they are on some occasion praiseworthy.

I have not provided anything close to an exhaustive picture of views of blame that might be combined with praise in the estimability sense. And even for the views I have discussed, further issues must be addressed before we have in place an adequate account of moral responsibility that includes estimability. For instance, do deserved praise and blame draw on the same notion of fairness? It is sometimes claimed that praise invokes a *comparative* notion of fairness, since one's being undeservedly praised does not obviously harm the person praised, and seems to be unfair only to the extent that it deprives *others* the praise they deserve (Watson 1996: 284; Nelkin 2009: 155). I am inclined to think that a distributive notion of fairness is common to deserved praise and blame, even if *other*

---

[50] The estimability theory may of course fail to specify a *complete* desert base for Uri's being deserving of praise. If, for example, there is a *historical condition* on responsibility, Uri's action will also have to have been performed from unsheddable values (and so not be e.g. performed only as a result of a benevolent neuroscientist's manipulating Uri's brain). But even if we should ultimately adopt a historical condition for interpersonal praise, it would be dialectically infelicitous to build separable conditions like this into a view of estimability, especially given my present aim of combatting the demand prejudice.

norms of fairness contribute to the explanation of the badness of an agent's being undeservedly blamed (e.g. a norm based in the principle that 'ought implies can' (Nelkin 2009: 161-2)). This idea fits naturally with a view of moral responsibility as an essentially interpersonal affair, a joint enterprise of managing a certain class of benefits and harms. A defense of this view, along with discussion of related issues, must however wait for another time. My aim in this section was only to give plausibility to the proposal that estimability praise is a subspecies of a broader type of responsibility that might include accountability blame. Responsibility in the deserved moral address sense is poised for this task, for it captures the key features of the accountability theorist's view of blame. Praise and blame in the attributability sense of responsibility can be contrasted with the more robust praise and blame of *responsibility in the deserved moral address sense* [R-DMAS]. Philosophers *not* antecedently committed to demand-based views of blame (e.g. those with "protest" views) have a shorter route to an appealing version of R-DMAS, for they are already free of the demand-prejudice afflicting an understanding of interpersonal praise.

5. Concluding Words

I hope to have shown that interpersonal praise cannot be understood in terms of deontic concepts like "moral demand", and that the assumption that it can be so understood—what I have called the *demand prejudice*— has a stronghold largely owing to the kind of dominance that accountability conceptions of responsibility have enjoyed. I have proposed an axiological (or evaluative) conception of responsibility, in the *estimability* sense, that characterizes our aspirational stance of holding agents up to interpersonal ideals. It is in

light of this latter stance, which I identify as one of normative hope, that our praise of others

is fruitfully understood.

4

Standing to Praise

0. Introduction

A widely held view in the moral responsibility literature is that although some agent may be blame*worthy*, it might nonetheless be inappropriate for *me* (or some other particular person) to blame that agent. That is, in virtue of one's relation to the blameworthy agent, one might lack the *standing* to appropriately feel and express attitudes of blame, like resentment or indignation. Various conditions have been identified as necessary for "standing to blame", of which the most commonly cited are perhaps *non-hypocrisy*, *non-complicity*, and *non-meddling* conditions.[1] The hypocrite, for instance, is sometimes taken to lack the standing to blame because in blaming another for a wrongful act that he himself refuses to 'own up to' (i.e. express remorse, or apologize, for), the hypocrite expresses wrongful disregard for the moral equality of persons (Wallace 2010: 329).

On the face of it— and as one would be inclined to conclude from a survey of the literature on moral responsibility— such "standing" conditions have no analogue in the

---

[1] See Cohen (2006); Radzik (2011); Scanlon (2008); Smith (2007); Wallace (2010); Wertheimer (1998).

domain of praise. While it is generally assumed that persons sometimes fail to be "*blamer-worthy*" (Friedman 2013: 273), similar restrictions seem not to apply to the '*praiser-worthiness*' of persons. We might attempt to ground this difference in other asymmetries between praise and blame, e.g. in blame's tendency to cause pain, or perhaps in its characteristic force or *sting* (Hieronymi 2004). Perhaps praise differs from blame in being like good company: 'the more the merrier'. Given the abundance of research on norms of blame, however, we should expect there to be *some* philosophical reflection on the idea that our responsibility practices display sensitivity to norms governing "standing to praise".[2] The absence of such reflection is partly symptomatic of the fact that our vocabulary praise is generally impoverished relative to that of blame.[3] But, if we hope to understand the ethics of our responses to a full range of morally responsible agents, it is incumbent upon us to enrich our vocabulary of praise. In this paper, I take a step towards this goal by making a case for the existence, and *importance,* of norms of standing to praise. Specifically, I argue that the praise of candidate praisers is governed by what I call the "familiarity condition": *for S to appropriately praise T for ϕ-ing, S's praise must be based in familiarity with the value expressed in T's ϕ-ing.*[4] Praise that fails to satisfy this condition

---

[2] Markovits (2012) is almost an exception. Although there is no substantive discussion of standing to praise in her paper, several of her examples invite such discussion, which Markovits (2012: 310) anticipates in stating, by way of conclusion, that she thinks "that facts about how we are willing to act play a role in answering the (normative) question concerning what kinds of acts we have standing to praise and blame."

[3] A further (non-vindicatory) aspect of the explanation is that talk of "standing" derives at least in part from legal discourse, where "standing" or *locus standi* refers to the right a plaintiff must have in order to bring suit to court (*US Const*. Art. III, sec. 2). For influential discussion of modern standing law, see Fletcher (1988).

[4] Although I will speak mostly of actions as the objects of praise, the condition can be modified to target attitudes or traits. (To be sure, attitudes and traits are often implicated in the praise of actions.) For an influential paper in favor of divorcing warranted blame (and praise) from voluntary control, see Adams (1985). Adams' focus is on responsibility for

is vulnerable to a distinctive kind of criticism. Reflection on the familiarity condition sheds light on a valuable feature of our responsibility practices that is otherwise overlooked, namely the phenomenon of co-valuation that successful praise engenders.

I proceed as follows. After clarifying the relevant sense of praise at issue (section 1), I introduce two relatively uncontroversial norms governing the praise of candidate praisers: an epistemic condition and a right-relationship condition (section 2). The function of section 2 is to motivate a shift in attention from the target of praise to the person doing the praising, i.e. the praiser. While the epistemic and right-relationship conditions can be extrapolated from existing work on responsibility, section 3 introduces a novel condition — the "familiarity condition" of standing to praise. The elucidation of this condition is this paper's main aim. In addition to descriptive aptness, a desideratum of my account is a moderate defense of the *appropriateness* of the familiarity condition. Accordingly, in section 4 I propose a route for justifying the familiarity condition, one based in considerations comparable to, but importantly distinct from, those underlying the inappropriateness of hypocritical blame.

1. Blame and its Contrary

"Praise", as it is understood in this paper, is the opposite of blame. In saying that "voluntary feelings and actions are praised and blamed" Aristotle (1109b30-1) seems to take praise and blame to form a contrasting pair. And, it is generally held that praiseworthiness and

---

attitudes (emotions and desires); see Sher (2001; 2006: 51-70), for the related view that blame can be deserved for *traits*. For other sophisticated defenses of non-volitional views of responsibility, see Scanlon (1998) and Smith (2005; 2008).

blameworthiness are opposed poles of moral responsibility. Although the everyday use of the word "praise" tends to be closely associated with overt acts or expressions—as Coates and Tognazzini (2013: 4-5) write, "the fit between being praised and having one's praises *sung* seems quite tight"— there is no contradiction in the idea of *private* praise. One can certainly *think highly of,* or admire, someone without manifesting one's attitudes in behavior. On a view rooted in Strawson (1962), praise and blame are paradigmatically instanced in "reactive attitudes" like gratitude and admiration, which can of course be had, or *felt*, privately. And indeed, the pleasure of being overtly praised depends in large part upon the presupposition that praiser has certain positive evaluative attitudes. While this paper does not presuppose a Strawsonian account of responsibility—more on this momentarily— owing to the view's prominence, the idea of an '*attitude* of praise' is far from being an unfamiliar one in moral philosophy. In any case, I will take praise to consist in express*ible* but not essentially expressed attitudes.

I remain neutral on exactly what *kinds* of attitudes are constitutive of praise, but will treat emotions like admiration and gratitude as paradigmatic praise attitudes. While it is sometimes held that emotions of this sort are "reactive attitudes" that can only be justified if we are sources of action in a causally demanding (and perhaps naturalistically unbelievable) sense—what Pereboom (2001; 2013; 2014) calls responsibility in the "basic desert"-entailing sense—one might deny (or be skeptical) that we are agents who can be morally responsible in this demanding sense *without* denying that emotions like gratitude and admiration are praise-manifesting attitudes that presuppose a more minimal, though legitimate, form of moral responsibility (e.g. responsibility in what is sometimes called the

"attributability" sense⁵).⁶ In what follows I treat attitudes like admiration and gratitude as paradigmatic attitudes of praise, but I remain neutral both on a) what *type* of responsibility—and so, what set of agential capacities—such attitudes presuppose, and on b) whether these emotions are necessary for praise.

Although I take praise to consist in (not necessarily expressed) attitudes, expressed praise will be central to this paper. This should not be surprising given that discussion of standing to blame heavily concerns expressions of blame. The offensiveness of the hypocrite's blame, after all, depends heavily on his *expressing* blame towards a person whom he lacks the standing to blame. Since norms of interpersonal praise are our topic, unless otherwise noted, in what follows "praise" refers to communicated expressions of praise-attitudes.⁷ While the *word* "praise" is sometimes used to refer to encouraging behaviors motivated solely by instrumental or consequence-driven considerations (e.g., positive reinforcement, habit-formation), this sort of behavior, directed to pets-in-training

---

⁵ To hold an agent attributability-responsible for an action is roughly to view that action as expressive of some excellence or defect of her character or practical identity. Praiseworthiness in the attributability sense is thus a non-consequentialist form of responsibility that one kind of responsibility skeptic can accept. See Watson (1996) for influential discussion of the distinction between attributability and accountability, and Shoemaker (2011; 2013) for development of the idea that moral responsibility admits of various types.

⁶ Indeed, it is Pereboom's (2013: 190) view that "denying basic desert praise- or creditworthiness would not jeopardize [axiological] assessments of goodness."

⁷ Expressed praised, strictly speaking, is not essentially *communicated*. A solitary trip to the cemetery might be expressive of one's admiration or gratitude. Still, for present purposes I restrict myself to praise that *is* communicated (but not necessarily via speech acts; gestures may be conventionally implicated to communicate praise). Expressed praise should further be distinguished from "reward", just as expressed blame should be distinguished from punishment. Since one can qualify for a reward without doing anything praiseworthy, one's being rewarded, even deservedly, does not imply either being praised or meriting praise.

and persons alike, is not my topic.[8] Expressed praise *is* behavior, but it is behavior expressive of attitudes that represent their target as being *worthy* of praise (i.e., praiseworthy).

Lastly, although my discussion is not restricted to praise considered as a *moral* phenomenon—intuitively, one can praise Wagner for the aesthetic range of the *Ring Cycle* despite disapproving of him morally—praise is nonetheless in my view responsive to the sensitivity of agents to norms and ideals (be they moral, aesthetic, epistemic, etc.).[9] For this reason, favorable attitudes towards 'non-agential' phenomena like mountains and sunsets fall outside of my purview (even if we sometimes use the word "praise" in these contexts). My argument, however, does not require acceptance of a picture of praise as unified across normative domains. To the extent that a distinction can be drawn between the nature and norms of praise, my interest here is in the latter (with the qualification that praise be understood as non-instrumental). For this reason, those who think that there is an essential difference between moral praise and praise in other normative domains—perhaps because moral praise is uniquely concerned with *motives*— are welcome to read the following in accordance with such commitments. Moral praise, in any case, looms large—even if not exclusively— in what follows.

---

[8] The point here is not that we cannot genuinely praise pets. Given my silence on the agential capacities presupposed by praise-attitudes, I have left open the possibility that non-human animals may sometimes be (appropriately) praised in the same sense that persons are praised. The point is only that *instrumental* 'praise', which is often used to habituate pets (and, especially young, persons), is not what I mean by "praise". Thanks to Guy Kahane for discussion of this point.

[9] While the objects of aesthetic appreciation are not limited to the agential, given my definition of 'praise', 'aesthetic praise' is here limited to *artistic* praise. Artistic praise might, as some have held, itself be understood as a species of praise of *creativity*, which is surely an agential phenomenon. See Kane 1996: 81; 89; 160-3 *passim*; Russell 2008.

2. Praiseworthy but not Appropriately Praised *by me*: Motivating the Idea

I turn now to two norms of praise that concern the prais*er* rather than the target of praise: an epistemic condition and a right-relationship condition. I am not concerned to argue that these should be understood as conditions of standing to praise (nor do I think they are). Rather, their function is to help pave the way towards reflection on such conditions (which are the topic of section 3).

First, the epistemic condition of praise. Suppose Robert is praiseworthy for some courageous deed. If I lack reason to believe that he performed a deed that was praiseworthy, intuitively I am not in the right relation to Robert to praise him, praiseworthy though he is. In other words, we recognize an **epistemic norm of praise [ENP]** according to which:

> It is inappropriate (absent special justification) for A to praise B for φ-ing if it is not reasonable for A to believe that B is praiseworthy for φ-ing, and if A's praise is not based in this belief.[10]

If I believe that Robert is praiseworthy on insufficient evidence, or on the basis of an unreliable or poorly functioning belief-forming mechanism [or…(*please insert your favorite criteria for 'A's reasonably believing that B for φ-ed'*)] then I might fail to satisfy ENP, and so, violate the epistemic norm of praise. Additionally, to satisfy ENP my belief about Robert and my praise of him must be suitably related. It is not enough that I merely *have* the relevant belief; the reasonableness of my praise depends on its being *based in*—

---

[10] My discussion and formulation of ENP is indebted to Justin Coates's (2016) discussion of an analogous norm on blame. The point about basing, however, appears to be lacking from his analysis.

and so, expressive of— the belief that Robert did a good thing. In other words, I do not satisfy ENP if the relevant belief about Robert is no part of the explanation of my praise.[11]

Now, in what does the inappropriateness of ENP-violating praise consist? One proposal would have it that, if I praise without satisfying ENP, I am liable to mislead others. Although Robert in this case is praiseworthy, since my praise of him is not rooted in a reasonable belief that he is praiseworthy, my praise will have the tendency to mislead others. Given, however, that we attach a special kind of importance to the praiseworthy and blameworthy features of persons, this is unlikely to be the full story. For it has nothing to do with the ethics of praise in particular. That is, *any* instance of asserting or implying some proposition "p" without reasonably believing "p" is apt to mislead.[12]

To see what a better proposal would look like, consider what Coates takes the badness of epistemically unjustified *blame* to be: "when we blame someone without satisfying [the epistemic norm of blame], we reveal ourselves to be insufficiently concerned with the value of innocence or with making sure that we blame only if the agent in question deserves to be blamed."[13] As formulated, this idea will not explain the inappropriateness of epistemically unjustified praise; in praising someone it is not normally that a *presumed*-guilty person is implicated as instead being blameless. (Praise is not exoneration.) But, a satisfactory explanation is in the offing. In praising I represent the object of my praise as being exemplary in some respect, and on this basis, as meriting

---

[11] I put aside for the moment questions surrounding *partial* basing. I return to these, however, in defending the familiarity condition, as it too includes a basing-requirement.
[12] In thanking, applauding, etc. it may be (other things being equal) conventionally implicated (though not asserted) that the target of these expressions is *praiseworthy*. Such conventional implication, I take it, ordinarily licenses inferences about praiseworthiness; cf. Coates 2016: 464.
[13] Coates 2016: 462

regard or esteem. Plausibly, then, praise that flouts ENP manifests insufficient concern for

the value of *having merited* high regard. Given that we do not praise people for just any

old actions, my praising some person without satisfying ENP reveals me to be

insufficiently sensitive to the value manifest in the conduct of individuals like Desmond

Doss, Paul Farmer, John Stuart Mill, and Rosa Parks, i.e. genuinely praiseworthy persons.[14]

In addition to an epistemic norm, we recognize a **right-relationship norm** of

praise. The relationships we have with praiseworthy agents affect the kind of praise it

---

[14] Reflection on ENP invites worries about a more radical kind of skeptic, one who insists that we *never* satisfy the epistemic condition of praise. The skeptic I have in mind does not deny that we can *be* morally responsible (and so, praiseworthy) in the actual world. Rather, she holds that in every particular case in which one might praise, one cannot rule out praise-undermining motives in the candidate target of praise. For the radical skeptic about praise, the opacity of motivation undermines epistemic warrant for praise in every case. This sort of skepticism might remind one of the skepticism about blame that Gideon Rosen (2004) defends on the grounds that it is extremely difficult to know that any particular bad action was done from akrasia—"the only possible locus of original responsibility [being] an akratic act"—rather than from non-culpable ignorance of relevant facts or norms (Rosen 2004: 307; 311). Although Rosen confines himself to skepticism about warranted blame, it is not difficult to see how this form of skepticism could be generalized to apply to praise. The radical praise skeptic might claim ignorance (or at least uncertainty) about the motives from which any prima facie exemplary action is performed. (Kant (1785) gives voice to this view in writing that it is "absolutely impossible to settle with complete certainty through experience whether there is even a single case in which the maxim of an otherwise dutiful action has rested solely on moral grounds and on the representation of one's duty" Ak 4:407.) *True*, praise is not harmful in the way that blame can be, but this sort of difference can only provide pragmatic reasons to be more confident in assessments of praise than blame. How might we respond to the radical praise skeptic? First, the standards for justified assessment of praiseworthiness are plausibly less demanding than she makes out. Indeed, the radical praise skeptic might be charged with changing the explanatory background against which I ask 'am I justified in judging praiseworthy this agent?' (see Björnsson and Persson 2012). If we assume that the practical stakes can affect the reasonableness of beliefs about praiseworthiness, then perhaps in *some* cases the epistemic bar might rest very high for justified praise (see Coates 2016: 469-71). Still, these will presumably be the exceptions. Moreover, even if the radical praise skeptic is right in thinking that ENP is never satisfied, she can nonetheless take onboard *further norms* governing praise. Assuming that the justification of praise comes in degrees, one's failure to satisfy *non*-epistemic conditions of praise might *further diminish* the already insufficient justificatory position one is in to praise.

would be appropriate to give them. For example, the fact that Shaun's parents are praiseworthy for their exemplary parenting does not, by itself, put me in the right relation to be *grateful* to Shaun's parents. To put this slightly stiltedly, Shaun is the direct beneficiary of his parents' good will, and for this reason he stands in the right relation to them for gratitude. If there is an appropriate attitude for *me* to have towards Shaun's parents, it will presumably be impersonal one, like admiration.[15] Similarly, even if Tina's exemplary piano recital merits praise, I might not stand in the right relation to Tina to appropriately be proud of her for her achievement (as Tina herself and her piano-teacher presumably will be). The right-relationship norm is not unique to praise; it intuitively governs blame too. If X wrongs Y, *I* might not be in the right position to resent X, though my indignation might nonetheless be called for. Note that the right-relationship norm is compatible with its being true, in every instance, that if some subject S is praiseworthy, each person is in a position to praise S *in one way or another*. For this reason, the right-relationship norm is not properly considered a condition of "standing". Still, since the right-relationship norm entails that S's being praiseworthy *underdetermines* what attitude of praise is appropriate for a candidate praiser, it places the focus on the person doing the praising.[16] In this way, it moves us towards a discussion of standing to praise, proper. In

---

[15] This is not to deny that *Shaun* can intelligibly admire his parents. Extending a distinction from Darwall (2012) between individual and representative authority, we can say that if Shaun evaluates his parents impersonally (relative to their meeting a standard of e.g. exemplary parenthood) he may—just like any other "representative" of the moral community—admire them. By contrast, Shaun's gratitude to his parents involves a personal appraisal of them, as *benefitting him*.

[16] Some philosophers distinguish this kind of "positionality" of assigning responsibility from "standing". According to Bell (2013)—who argues that we should abandon commitment to standing to blame— blame is positional, where this means that the *fittingness* conditions on blame attitudes (and not just norms on their *expression*) are (sometimes, but notably for her, not usually) relationship-dependent.

what follows, unless otherwise noted, it can be assumed that ENP and the right-relationship condition are satisfied.

3. Familiarity and Standing

We recognize a norm that reflects the importance of the praiser's familiarity with the object of her praise. Consider the feelings of alienation that war veterans report feeling upon receiving the praise of individuals with little grasp of the realities of war. "*Thank you for your service*," Hunter Garth, a Marine veteran who served and almost died in Afghanistan, said, "is almost the equivalent of "I haven't thought about any of this.""[17] If aversion to this sort of praise were reflective of an idiosyncrasy of Garth's, then attempting to draw from it general insights about the ethics of praise would be misguided. But, Garth is not alone in feeling uncomfortable about the praise of strangers to war: "[t]o some recent vets — by no stretch all of them — the thanks comes across as shallow, disconnected, a reflexive offering from people who, while meaning well, have no clue what soldiers did over there or what motivated them to go." [18] If these veterans universally failed to view themselves as praiseworthy, then "standing" would not provide the best explanation for the uneasiness they experience in being praised. To be sure, it is well documented that veterans sometimes struggle with (sometimes severe) feelings of guilt about aspects of their service (especially aspects including their causal contribution to the deaths of fellow combatants).[19] While some of these veterans may judge themselves to be praiseworthy for *no aspect* of their

---

[17] Richel (2015)
[18] Richel (2015); see also Sherman (2015, esp. 25-41); Carter (2011).
[19] See Sherman (2010: 89-111).

service (perhaps judging their very participation in the war as morally impermissible), the reality is often more complex, with localized judgments of guilt sitting alongside other judgments of praiseworthiness (e.g. for one's contribution to a particular military effort). Granting that the psychological profiles of veterans are bound to be varied and complex, the veterans I have in mind are not ones who view themselves as unfitting targets of praise for their service. And, these are not make-believe veterans. Garth, for instance, "appreciates thanks from someone who makes an effort to invest in the relationship and experience. Or a fellow vet who gets it."[20]

I propose that owing to one's lack of *familiarity* with the value expressed in another's praiseworthy action, one lacks a certain kind of standing to praise. The familiarity condition is more demanding than ENP, the above epistemic condition. One can know (or justifiably believe) via testimony why one is praiseworthy, *without thereby* being familiar with those reasons. It is not as if non-veterans are ignorant of what is involved in military service. Still, Garth reports that "when he gets thanked it can feel self-serving for the

---

[20] Richel (2015) Further complexities are raised to the extent that we doubt that Garth in fact *is* praiseworthy—and one's answer to this question will depend at least in part on one's view of the morality of the war in question—but let us suppose that there is something praiseworthy about Garth's service. Or, if necessary, replace Garth with a war veteran whose service was, in your view praiseworthy. Consider Desmond Doss. Doss, a conscientious objector and combat medic who (on religious grounds) refused to bear arms (or otherwise take enemy lives), nearly singlehandedly rescued over fifty wounded and stranded soldiers at the Battle of Okinawa. (He also strove to rescue several wounded enemy combatants (although they did not survive)). According his Medal of Honor citation he received the award for "conspicuous gallantry and intrepidity in action above and beyond the call of duty…Pfc [Private First Class] Doss refused to seek cover and remained in the fire-swept area with the many stricken, carrying all 75 casualties one-by-one to the edge of the escarpment and there lowering them on a [self-devised] rope-supported litter [i.e., stretcher] down the face of a [400 foot-high] cliff to friendly hands." For the entirety of Doss' Medal of Honor citation, see: https://catalog.archives.gov/id/595188. Cf. http://www.nytimes.com/2006/03/25/us/desmond-t-doss-87-heroic-war-objector-dies.html

thankers…[suggesting] that they somehow understand the sacrifice, night terrors, feelings of loss and bewilderment. Or don't think about it at all."[21] In spite of knowing what war's about, they fail to "understand" something. I take it that the relevant sense of "understanding" here is one that might also be used in efforts to comfort a mourning person with words like: "I understand what you're going through". This expression does not mean 'I can explain why you are in pain', but something closer to, 'I am familiar with your pain, having suffered a similar loss.'

Might familiarity be a *kind* of knowledge? Possibly. But, since those who praise Garth do *know* that he is praiseworthy, to say only that familiarity is, or reduces to, knowledge is to shed little light on what is special about the phenomenon in question. For the time being, I use "familiar with" as a placeholder for a kind of relation that will become elucidated via considering several examples. At this stage, certain ways of characterizing familiarity might suggest themselves, e.g. knowledge by acquaintance (rather than description); emotional understanding; understanding what/that. I will not prejudge the matter by assuming which (if any) of these might be identified with familiarity.

Instead, let us consider another example. A literature student named Pete might have testimonial knowledge that Proust is a great writer (suppose his literature teacher told him so). And yet, Pete's admiration of Proust's literary excellence would be untoward in the absence of familiarity with Proust's work. In this kind of case, familiarity will typically be acquired through *reading* Proust (where this involves a range of cognitive and imaginative activities).[22] Admittedly, a constraint on the intelligibility of Pete's admiration

---

[21] Richel (2015)

[22] *How much* of Proust suffices for *familiarity*? Must one read him in *French* for familiarity? Here I defer to Proust scholars, disagreement among them notwithstanding.

might be that he be familiar with *some kind* of literacy excellence. Still, having never read (or listened to audio recordings of, etc.) Proust, he intuitively lacks familiarity with Proust's literary excellence.

It may be tempting to suppose that Pete's praise must be based in some ulterior motive, like the desire to gain favor. This desire could take various forms. Perhaps Pete wants to impress others with his literary tastes, tastes to which he cannot truly lay claim. Or, if Pete is Proust's contemporary, we might imagine him praising *In Search of Lost Time* with the desire to win Proust's favor. In the latter case, Pete's praise will constitute an instance of *flattery* (Eylon and Heyd 2008).[23] Now, if Pete praises because he wants to impress others, because he wants to win Proust's favor, or from some other ulterior motive, it will be no mystery why his praise is objectionable. This sort of *ulteriorly motivated praise*, as we might call it, would be intuitively objectionable in virtue of its being manipulative or surreptitiously deceitful. This proposal, however, fails to get to the bottom of what's objectionable about unfamiliar praise. For there is no reason to suppose that Pete must have an ulterior motive. To be sure, in addition to being objectionable owing to unfamiliarity, one's praise might be manifested with a desire to gain favor. If this were true of Pete, his praise would be doubly objectionable, for he would be both a) unfamiliar with the value of that which he praises, and b) his praise would be ulteriorly motivated. But we can stipulate that Pete's praise is not based in the desire to gain favor, or in any other ulterior motive. The alienation felt by veterans, after all, does not seem to depend on the assumption that those who praise them without familiarity do so with ulterior motives. "For

---

[23] That is, we can suppose it "makes use of excessive commendatory language in describing the qualities or record of another person for the purpose of creating a favorable attitude in that person towards the flatterer" (Eylon and Heyd 2008: 686).

most of us, I suspect," writes Richel, "offering thanks reflects genuine appreciation — even if ill-defined."[24]

One might think that even if it is not ulteriorly motivated, the problem with Pete's praise is *really* that it conversationally implies that Pete has done something he has not, i.e. read Proust's work. While Pete might lack the desire to deceive, his praise *itself* is somehow deceptive, owing to the meaning of its terms or the pragmatics of the assertion. This explanation locates the inappropriateness of the praise in the conversational pragmatics of praise, or some other linguistic feature of the utterances of praise. This proposal is worth exploring. Suppose Pete expresses his admiration of Proust's *In Search of Lost Time* by saying

[S] 'The novel is exceedingly beautiful.'

What explains the inappropriateness of Pete's asserting [S1]? Intuitively, it is the fact that [S] implies that the speaker has read the novel. That is, [S1]—regardless and independent of any desire Pete might have to deceive—seems to license the inference that the speaker is acquainted with the object of his praise, which in this case, would naturally require that he has read (or listened to an audiobook, etc.) of the novel.[25] But what is it about the statement [S] that implies that its speaker has read the novel? It is presumably something about its use of evaluative terms. After all, the statement 'the novel is divided into seven volumes' does not imply that's its speaker has read the novel. The "acquaintance inference"

[24] Richel (2015)
[25] Or at least some representative portion of it.

refers to the phenomenon wherein evaluative assertions involving aesthetic or taste predicates seem to imply that their speaker has first-hand experience with the objection of evaluation (Ninan 2014; Franzén 2018). The sort of acquaintance implied will depend upon the objection of evaluation; while the statement 'Milan Cathedral is stunning' implies its speaker has *seen* the cathedral (or perhaps a photo or video of it), [S] implies that the speaker has *read* (or perhaps listened to an audiobook of) the novel (Franzén 2018: 2). Pete, of course, has *not* read the book. So, in praising the novel by saying [S1] Pete's assertion suggests that he is acquainted with something he is not, namely the contents of the novel.

Can reflection on the acquaintance inference help explain what is objectionable about praising without familiarity? I do not think so. I grant that [S] implies that it's speaker, Pete, has read the novel. For this reason, in asserting [S] Pete will have flouted the linguistic norm [N], governing assertion that license the acquaintance inference. Consider that the way in which [N] governs [S] is also the way that [N] governs assertions of personal taste, like the following:

[L] 'The lobster rolls are flavorless'

[L] will be infelicitous if uttered by someone who has not has *tasted* the lobster rolls, for [L] licenses the inference that its speaker *has* tasted the lobster rolls. Furthermore, [L] seems to license the acquaintance inference in the same way that [S] does, namely by ascribing an evaluative predicate to some object. While there is disagreement about what

norm should be identified with [N],[26] it is agreed that [N]'s function is to explain what is *linguistically* infelicitous about a certain range of evaluative assertions, which range includes assertions of personal taste alongside aesthetic assertions. [N], the norm governing the acquaintance inference is for this reason unlikely to provide us with the kind of explanation we are looking for, namely one that sheds light on the ethics praise.

It is an incidental feature of Pete's praise that it violates [N]. To get to the core of what objectionable about unfamiliar praise let's turn to another case, one of moral praise. Suppose Lorenz admires and praises his neighbor, a retired lawyer named Adrian, for Adrian's generosity in volunteering several days per week to provide legal resources and advice to refugees seeking asylum, and to assist low-income non-English speaking immigrants with naturalization work. Now, in praising Adrian for his pro bono legal work, Lorenz's praise does not seem to license the inference that *he* does similar work. His praise will imply that he *knows* what it is to perform generous actions, and what it is about the status of the persons he's helping that renders his activity generous. But we can grant that he does know this. Suppose also that Lorenz lacks the motivation to help others in a manner that's analogous to Adrian's activity, or if he has *some* motivation in this direction, it

---

[26] The acquaintance inference, unlike conversational implicature, is not *cancelable* (Grice 1975; Ninan 2014: 297). That is, conversational implicatures to the effect that *p* can be felicitously cancelled by the speaker's tacking onto its end, *not p* (Grice 1975: 44).[26] For example, if I ask you, whether you plan on going to the festival on Friday and you *are* planning on going, but reply [W] 'I work on Friday', your utterance will license the conversational implicature that you are *not* planning on going to the festival. But, if you add 'but I will come afterward' to the end of [W] the implicature will be cancelled, and you will have violated no norm governing conversational implicature. The acquaintance inference—that licensed by Pete's praise [S1]— by contrast, is not cancellable. This can be illustrated by supposing that Pete's says,
[S2] 'The novel is exceedingly beautiful, but I haven't read it.'

always gets outweighed by others motives such that Lorenz never finds himself in fact helping those who are, in one way or another, in genuine need. His career provides him with various opportunities to mentor others, but he fails to pursue these or any other initiatives aimed at improving the lives of those who are professionally, socially, or economically 'below' him. It's not that he fails to believe that does so would be a *good thing*, but generous actions like these beliefs fail to engage his emotions, motivations, and other beliefs in a strong and consistent enough manner for him to ever pursue, or resolve to pursue, them.

Like Pete's, Lorenz's praise and admiration are not ulteriorly motivated. When Adrian comes home from volunteering he and Lorenz occasionally chat, and Lorenz is filled with admiration for Adrian, asking about the details of the plight facing his clients and advisees. Adrian is not put-off by Lorenz's praise in the way that Pete's interlocutor might be, for Lorenz's statement 'wow, that's a really laudable way to spend a Wednesday!' does not imply that he has done something that he has not. Adrian is of the opinion that Lorenz is an acquisitive and career-driven man, and while he generally enjoys his chats with Lorenz and takes Lorenz to be sincere, witty, and well-meaning, he occasionally feels that Lorenz exoticizes Adrian's volunteer work. He can't put his finger on it, but there's something that Lorenz sometimes fails to *get* about challenges facing the people he spends much of his time with.

Sincere and non-deceptive though it may be, if it is not grounded in familiarity with the relevant values, one's praise is liable to be unwittingly hollow. I assuage below the worry (at this point reasonable) that I am being too hard on well-meaning praisers. For now, let me introduce the

familiarity condition of standing to praise:

For S to appropriately praise T for ɸ-ing, S's praise must be sufficiently based in familiarity with the value expressed in T's ɸ-ing.

The basing condition matters for the following reason. Returning to Pete after years of admiring and praising Proust without familiarity, Pete *finally* reads *In Search of Lost Time* (or however much of it is necessary to meet the threshold for familiarity). Although he is now familiar with Proust's distinctive literary excellence (assuming, again, that Pete reads attentively and is not imaginatively impoverished, etc.), he might nonetheless continue to praise Proust *because* of his teacher's recommendation. That is, Pete's praise (after reading Proust) might not be based in, and expressive of, his newfound familiarity with Proust's excellence. In this case, Pete does not satisfy the familiarity condition.[27]

Is the "familiarity condition" a norm of *attitudes*, or does it govern *expressions* of praise-attitudes? With some qualification: *both*. In the first place, it governs attitudes (as well as dispositions). The objectionableness of Lorenz's praise is naturally located in the

---

[27] What if Pete's praise is based *partly* in familiarity with Proust work, and partly in the recommendation he receives from his teacher? Intuitively partial basing of this sort is compatible with the satisfaction of the familiarity condition. I suspect many, if not most, instances of praise will be partly based in factors beyond the value of the good-making features of the objects of praise. The recommendations of those we trust and respect are likely to remain as partial bases of our praise long after we have acquired familiarity with the things themselves. Whether this impugns one's praise is likely to be settled on a case-by-case basis, but the following counterfactual analysis for "sufficient basing" seems to point in the right direction: for S's praise to be sufficiently based in familiarity with the value expressed in T's ɸ-ing (e.g. Proust's authoring *In Search of Lost Time*), it must be the case that a) S's praise is based *predominantly* in S's familiarity with this value, and b) if S's praise has other, *external*, bases (e.g. another's recommendation), were these external bases for S' praise to be eliminated, S would nonetheless continue to praise (to a comparable degree) T in a way that is based in familiarity with the value expressed T's for ɸ-ing.

mismatch between what his attitude of admiration commits him to, valuationally, and what he *himself* can properly be understood as valuing. I will come back to this thought. Let me now clarify why I say the familiarity condition governs both attitudes and expressions. It is typically thought that the objectionableness of the hypocrite's blame is not confined to its being communicated. While it is presumably *more objectionable* for the hypocritical blamer not only to feel resentment, but to express it, the inappropriateness of his expression of blame has its roots in the lack of the hypocrite's standing to blame even 'in his own heart' the person he blames.[28] This consideration does not establish that conditions governing standing to praise *must* function similarly—after all, praise and blame admit of important asymmetries, but owing to the partial overlap between the objectionableness of hypocrical blame, on the one hand, and unfamiliar praise, on the other, it should not be suprising of norms of standing govern, each in their own way, attitudes and expressions of both blame and praise.

The idea that, absent familiarity, one lacks one kind of standing to have praise-attitudes will be implausible on cognitive theories of praise. That is, if to praise S just is to believe or judge that S is in some way good or exemplary, then lack of familiarity could hardly render praise inappropriate. Indeed, we charge with epistemic irrationality the person who has strong evidence for some proposition and yet fails to believe it. But, given that I can have strong evidence for the proposition 'that S is praiseworthy' (e.g. via expert testimony) *without* thereby having any familiarity with the value expressed in S's praiseworthy action, it would seem that on a cognitive theory of praise, familiarity is irrelevant to the appropriateness of one's attitudes of praise.

---

[28] See Wallace (2010: 324).

I grant that cognitive theories of praise have no place for the familiarity condition. Should this bother us (or rather, me)? It should not. Note first that cognitivist views of praise and blame are quite unpopular, even among those who accept rather minimal, attributability, views of what it is to hold another responsible.[29] Moreover, the conditions on standing to blame are *also* obviated on cognitive theories of blame. That is, whatever is being denied the hypocrite it is not the appropriateness of his judging or believing that another acted badly or wrongly. Similarly, my being complicit in your wrongdoing can hardly affect the rational basis of my judging or believing that you did something bad. And to return to Pete and Lorenz, despite their lack of familiarity with the relevant values, there is nothing objectionable about their *judging,* respectively, that Proust wrote a beautiful novel, or that Adrian's generous volunteer work is admirable Since, however, my aim in this paper is primarily to cast doubt on the assumption that there exists an asymmetry between blame and praise vis-à-vis the existence of conditions of standing, it is no concern to me that some theories of praise and blame (namely cognitive theories) have no place for standing conditions.

Let me turn to the objection that the "familiarity condition" is too harsh or demanding. Does it entail that I should never praise another if I lack familiarity with the

---

[29] Zimmerman is thus an outlier among proponents of responsibility-as-attributability. On his view, "[p]raising someone" is a matter of "judging that there is a "credit" in his "ledger of life," or a "positive mark" in his "report-card of life," or a luster on his "record as a person"; that his record has been "burnished"; that his "moral standing" has been "enhanced"" (Zimmerman 1988: 36). By contrast, see Arpaly and Schroeder (2014: 160; 161), who write: "[t]o grasp that someone is blameworthy for an act is not yet to blame the person... One must also be *against* what's wrong or bad, in a way that Satan might not be, in spite of understanding that one's action was done out of bad will." See also Sher (2006: 75), whose project is formulated as answering "what blaming someone adds to believing that he acted badly".

value expressed by her praiseworthy action? Fortunately, it does not. One's failing to satisfy the familiarity condition does not always give one *decisive* reason to withhold from praise. Or, more accurately, since attitudes of praise are not under voluntary control (and so, not the sort of thing one *withholds*), one's failing to satisfy the familiarity condition is not always, on balance, *criticizable*. To see why this is so, contrast the above Proust-praiser—who, by stipulation fails to aspire toward familiarity with Proust's work (he's satisfied with testimonial knowledge of Proust's greatness)— with a budding student of opera who praises Verdi. This Verdi-praiser is unable to appreciate the quality of Verdi's work to the extent possible for the opera aficionado, but she aspires to see what the aficionado sees. Does the familiarity condition issue the verdict that the Verdi-praiser's praise is objectionable? By itself, *yes*. But the familiarity condition is not the only normative consideration relevant. Her praise is objectionable in that the Verdi-praiser lacks familiarity with the object of her praise. But, this is only one dimension along which her praise can be evaluated. Since this opera student aspires towards familiarity with Verdi's work, her praise might be part of a genuine effort to *come to value* Verdi's work in the way that the aficionado does. As Agnes Callard puts it, it is "a fact of life that we act not only *from*, but also, at other times, *for the sake of acquiring*, knowledge of value."[30] For this reason, even if the aspirant's lack of standing gives her a pro tanto reason *not* to praise, she might possess an all-things-considered reason to praise. In this case she will not be criticizable, on balance, for her unfamiliar praise.[31]

---

[30] Callard (2016: 131). I am agnostic, however, about whether increased familiarity is a matter of *knowledge* acquisition.

[31] There are several ways to understand the claim that one is not, on balance, criticizable for x. One is that one set of considerations (C1) renders the agent commendable for x and another set considerations (C2) renders one criticizable for x, and that C1 is stronger than

A related consideration that helps with the 'demandingness worry' is that unfamiliarity intuitively comes in degrees. Even if we accept some threshold of familiarity requisite for standing, there is no obstacle to supposing that this threshold is flanked by degrees of familiarity and unfamiliarity. Familiarity, in other words, need not be construed as all-or-nothing. And, I submit, it *should not* be construed as all-or-nothing. Two people might both have read a sufficient amount of Proust to count as familiar with Proust's literary excellence, although one might be *more familiar* than the other. The same point naturally holds of unfamiliarity; two people can be unfamiliar with Verdi, though one unfamiliar to a greater degree than the other, perhaps never even having heard of the composer. Although the aspirant does praise without a certain sort of standing (that provided by familiarity), her standing to praise becomes less and less diminished as she praises (on the assumption, that is, that her praise is itself part of a process that contributes to her being less and less *un*familiar with the value of the object of her praise). The familiarity condition thus helps us make sense of the clumsiness with which aspirants and other novices praise, while granting that they may have overall reason to do so.

One might worry that I have made things too easy for myself by focusing on aesthetic cases. The opponent who raises the demandingness worry might urge me to return to Garth, the war veteran. Suppose one were to praise a praiseworthy veteran in the following, self-consciously unfamiliar, way: *'I can't understand the sacrifice and courage involved, but I want you to know that I am grateful for your service.'* Call this an instance

---

C2. On this construal of on balance non-criticizable, C2 functions a bit like an excuse. Alternatively, one might understand the claim that one is not, on balance, criticizable for x to mean that the reasons for criticizing are extinguished by the reason for commending, such that the latter is justified.

of *self-consciously unfamiliar* praise (or, sᴄᴜ praise). Is it really my view that one has (even defeasible) reason to withhold from sᴄᴜ praise if she lacks familiarity with the realities of war? The first thing to note is that one might be familiar with the value expressed by the target of one's praise—say, the value of courage—without having had first-hand experiences of warfare. It does not follow from the fact that the praiser is unfamiliar with the value of courage in battle that the praiser lacks familiarity with the value of courage *in other domains*. An individual might be familiar with *one* determination of some value V without being familiar with other determinations of V. We often come to appreciate (or, at least *approximate* appreciation of) the value of another's achievement by comparing it to achievements with which we are familiar. This kind of response, however, can only get us so far. sᴄᴜ praise, after all, is naturally understood as praise in which the praiser genuinely *is unfamiliar* with the relevant value. That is— without taking a stand on whether familiarity with courage in other domains can be sufficient for familiarity with courage in battle— we can assume that the sᴄᴜ praiser in fact lacks familiarity with the relevant value manifested by the target of her praise.

Now, to respond to my objector: *yes*, I maintain that there is something defective about sᴄᴜ praise. But, note that in making explicit one's lack of familiarity, the person offering sᴄᴜ praise seems to acknowledge the fact of this defectiveness. To bring out this feature of sᴄᴜ praise, compare it with 'self-consciously hypocritical blame' [sᴄʜ blame]: 'I know I'm always remorselessly late for our meetings, but I'm really angry at you for being late!' To be sure, the self-conscious aspects of both sᴄᴜ praise and sᴄʜ blame *speak well* of the respective praisers and blamers. But to the extent that sᴄʜ blame *is* nonetheless *hypocritical* blame—and not an expression of someone 'owning up' to their past

wrongdoing— it remains partly objectionable. scu praise is intuitively less objectionable that scH blame —an intuition I vindicate in section 4—but if we thought there was something defective about unfamiliar praise to begin with, it is hard to see how the self-consciousness of one's unfamiliarity could *cancel* this defect. I propose, rather, that we understand the self-conscious aspect of scu praise as potentially *outweighing* the badness of the unfamiliarity of scu praise.

The criteria for "familiarity", as the above examples imply, depend upon the focus and depth of the praise in question. There is one variety of praise, however, for which only the expert has familiarity. This is a form of 'stock-taking' praise that we can call *critical praise*, for paradigmatic instances of it are found in the art critic's evaluations. Take J. Hoberman's evaluation of the film directors Jean-Pierre and Luc Dardenne. They are hailed as "worker-priests of European art cinema…[T]he Belgian duo have perfected a sort of spiritually-infused social realism. Their movies are Robert Bresson reimagined as cinema verité."[32] The standing for critical praise seems to require a level of familiarity constitutive of expertise. The avid reader and Professor K may both have read enough Proust to count as *generally familiar* with Proust's value, but the value with which Professor K is familiar might allow for far more determinate description, e.g. of the value constitutive of Proust's ingenuity in the use of metaphor in the French language. This point helps characterize why the high school student who has read most of Proust and who thereby (we can stipulate) meets the lower threshold for familiarity, might nonetheless fail to be in a position for

---

[32] Hoberman (2012)

critical praise, as might be evidenced by his amateur critical praise at an academic conference of leading Proust scholars.[33]

The above example of the film critics' praise suggests a point that should be made explicit, namely that "familiarity" does not (generally) require that the praiser possess either the relevant capacities or kinds of experiences enjoyed by the praisee. These can be lacking in evaluatively familiar praiser. Intuitively, Hoberman's standing to praise the Dardenne brother's does not depend on his having comparable directorial experience or ability. This point about familiarity is not specific to critical praise. Stan might have the standing to praise his favorite soccer (/football) player (for the latter's athletic excellence) even if Stan has never played at a level even approaching that displayed by FIFA athletes. Standards for familiarity display domain-specificity. In Stan's case, satisfying the familiarity condition might require knowing how this player's stats compare with others', knowing how the player's qualities complement those of his teammates, perhaps being able to appreciate what is stylistically significant about his game-play, etc. Recall also Garth's statement that he appreciates thanks not only from "a fellow vet who gets it" but also "from someone who makes an effort to invest in the relationship and experience."[34] That a fellow vet can be one who fails to 'get it' suggests that shared wartime experience is insufficient

---

[33] This is especially so given that critical praise typically requires familiarity both with the *primary* object of criticism (in this case, the value of Proust's work), and with the secondary objects of criticism (the value of the work of those *evaluating* Proust's work). Illustration of this point is found in the following passage: "we can admire the artistry with which, as Gérard Genette brilliantly observes (1972: 42), Proust places a flowering plant in the Guermantes' courtyard just so that Marcel can "discover" it and use it as an image for the miraculous encounter between Charlus and Jupien. In all such cases, elements of Marcel's reality, dutifully reported in *My Life*, are, in the *Recherche*, artistic choices made by Proust." (Landy 2004: 44)
[34] Richel (2015)

for familiarity, and that a non-vet can appropriately praise suggests that such experience

is unnecessary

4. Justifying the Familiarity Condition

Is my discussion of the familiarity condition meant normatively or merely descriptively?

In elucidating a feature of our practices of praise, one does not commit oneself thereby to

a view regarding the appropriateness of that feature. For this reason, the reader may incline

toward an interpretation of the familiarity condition as simply describing a feature of our

responsibility practices. What motivated our inquiry, of course, was the sense that

something was *off* in cases of unfamiliar praise, like those involving the veteran and Proust.

But, what justification could there be for the familiarity condition? While providing an

answer to this question is not this paper's chief aim, let me venture a brief and tentative

reply. We might obtain an elucidation of the familiarity condition's normative significance

by attending to the function of praise, i.e. the role that praise occupies within our

responsibility practices. Consider the following, minimal, gloss on what is plausibly the

function of praise: the function of praise is for members of the moral community to affirm

shared values, or—isolating ourselves to the relation between praiser and praisee— for the

values of the praiser and praisee to give rise to *co-valuation*, i.e. a relationship constituted

by the value binding praiser and praisee.[35] An essential component of, and precondition for,

---

[35] This is not to say that we praise *with the intention of* displaying what we value. Praise is not "moral grandstanding", where "one grandstands when one makes a contribution to public moral discourse that aims to convince others that one is "morally respectable"" (Tosi and Warmke 2016: 199). I do not deny, however, that some instances of praise *can*

joint valuation of this sort is that one's praise of another generates a *resonance* of value. What is it for one's praise of X to give rise to resonance of value? Importantly, it is not enough that the praiser *judge* X valuable. Rather, she must *value* it (for its own sake), which importantly includes caring about X, which itself involves the disposition to feel joy and frustration (in myriad ways) depending on whether X is promoted or undermined.[36] In addition to emotional vulnerability relative to X, and the belief that X is valuable or good (in some way), valuing plausibly has "deliberative significance", which is to say that it involves a "disposition to treat certain kinds of X-related considerations as reasons for action in relevant deliberative contexts" (Scheffler 2010: 29). Now, for my praise of a praiseworthy agent to fulfil its function, my praise must be expressive of *my valuing* what the praisee's action revealed her to value. In the paradigmatic case of praise, I am moved to communicate to the praisee that I value her commitment to some value and my praise receives uptake in the praisee's response. Successfully expressed praise thus makes possible a relation in virtue of which praiser and praisee are capable of valuing *together,* or "co-valuing", which can be understood as "the valuational counterpart to shared agency" (Callard 2018: 130). Just as some actions cannot be performed alone, so too it is natural to assume that some valuable things (and aspects of what is valuable) cannot be valued in isolation. Co-valuing therefore is the relationship generated by the mutual acknowledgment of value resonance. While value resonance is a mental relation (the matching of the value represented in the attitude of praise with the value manifested in the

---

constitute grandstanding. Indeed, the Proust-praiser's praise intuitively reads as a case of *aesthetic* grandstanding (which would suggest that grandstanding—or something very much like it— characterizes not only to efforts to appear *morally* respectable but aesthetically respectable (and perhaps in other cases, epistemically respectable) too).

[36] See Anderson (1993); Kolodny (2003); Scheffler (1997; 2010).

praiseworthy action), co-valuing is the social phenomenon that arises when value resonance is mutually acknowledged.

Of course, not all instances of praise will display resonance of value. And for this reason, not all instance of praise will fulfil the function of praise: co-valuing. Things can go wrong in variety of ways. For example, praise can be directed at an action that is *not* praiseworthy. If the target of praise fails to exhibit the value it is represented as manifesting, we do not get resonance of value (and so, the expression of this praise cannot generate co-valuing). This kind of failure is not my focus. A more complicated way things can go wrong is illustrated by appeal to the familiarity condition. In cases of unfamiliar praise, there is an absence of value resonance in virtue of a feature of the *praiser*. The praiser who lacks familiarity fails to value the value (i.e., the valuable features e.g. the distinctive aesthetic qualities of Proust, the generosity of helping those in need) expressed by the object of his praise. This is because valuing X, unlike *judging* X valuable, is partly *constituted* by familiarity with X.

It will be helpful to return to an earlier example. Suppose that "Rob" is someone who praises Garth without any familiarity. Although Rob's praise is sincere and based in justified beliefs about Garth's praiseworthiness, suppose that Rob nonetheless fails to support (either by donating time, money, or any affective or attentional resources to) any initiatives or organizations aimed at improving the quality of life of veterans and soldiers. Although these omissions are symptoms of Rob's lack of familiarity with the value of military courage (and not constitutive of it), we can suppose that Rob lacks the disposition to give deliberative significance to reasons concerning veteran well-being.

If the function of praise is to co-valuation, Rob's praise will not fulfil its function. For, in contrast to judging something valuable, part of valuing something is to see that thing as providing one with non-instrumental reasons for action. Now, if Rob fails to value that for which he praises Garth, one might worry whether Rob is even capable of praising Garth. But unless we hold that it is constitutive of praise that it fulfils its function (which I see no reason to do), there is no barrier to supposing that Rob's is a genuine instance of praise. In spite of its being a genuine instance of praise, one that is furthermore epistemically justified and sincere, Rob's praise lacks a mooring in what he *values*. Owing to this lack of mooring (i.e. his lack of familiarity with the relevant value manifested in Garth's actions), Rob's praise is normatively defective. Since facts about his lack of familiarity explain why the quality of his praise precludes the resonance of value that is a component of successful praise, the expression of Rob's praise is unwittingly hollow. In other words, he lacks the standing to praise. Rob might *judge valuable* the good-making features of Garth's actions, but this is not enough for his praise to contribute to a resonance of values.

The defectiveness of Rob's praise does not consist in, or require, its being expressed. The veteran who recognizes Rob's praise as evaluatively unfamiliar may experience alienation in face of Rob's praise, and this alienation might be a way of registering that Rob's praise is hollow of the interpersonal meaning that it purports to bear. But this alienation is unnecessary for a failure of evaluative resonance, and the hollow expression of praise that gives rise to it is not *constitutive* value resonance.[37] Part of the

---

[37] By analogy: one might praise another with satisfying the epistemic condition of praise. One's praise will violate this norm regardless of whether it is expressed, and if it *is* expressed, regardless of the whether the praisee registering the perhaps uncomfortable truth that her praise comes from someone lacking sufficient evidence for taking him to be worthy of praise.

value of value resonance is that it enables agents to jointly value. Jointly valuing or co-valuing requires the expression of value, for the relationship depends upon these values being covertly shared. But the essential *precondition* for jointly valuing—the value resonance constituted by a match between that which the praiser values (in admiring or feeling grateful) and that which the praisee values—need not be expressed or communicated. Lorenz's moral admiration of Adrian's generous action, all by itself, precludes co-valuation. It is does this *by* precluding value resonance. Depending on how Lorenz expresses his praise, he might conversationally imply that he is familiar with the value of Adrian's action. In this case he will have flouted the linguistic norm that governs utterances licensing the acquaintance inference. But, as I have argued above, this norm is distinct from the familiarity condition.

While it would stretch the ordinary meaning of "hypocrisy" to call 'unfamiliar praise' "hypocritical", the non-hypocrisy and familiarity conditions both commend coherence between what one values on the one hand and one's other evaluative attitudes and conduct, on the other. The corresponding injunction here is to 'practice what you preach' (where 'practicing' includes one's affective and imaginative responses). It is for this reason that an agent's lack of standing to praise or blame is often evinced in his or her broader patterns of thinking, feeling, and acting. The foregoing discussion, however, allows us to see why unfamiliar praise and hypocritical blame are objectionable in different ways.

In blaming an agent for a wrong that she fails to hold herself to account for, the hypocritical blamer treats herself as a kind of moral superior. Hypocritical blame is for this

reason intuitively blameworthy. In praising without standing, by contrast, one does not treat oneself as morally superior. This is not to deny that unfamiliar praise is objectionable. After all, in praising without standing I praise in a way that deprives the object of my praise the involvement in an instance of co-valuation or value resonance. On the assumption that praise paradigmatically gives rise to mutually acknowledged value resonance, which itself is intuitively good, unfamiliar praise turns out to be a shadow of what it could be. It for this reason can understandably alienate the person praised. The proposal on offer is able to shed light on why praising (in contrast to blaming) without standing is not obviously blameworthy. Although unfamiliar praise is normatively defective, and for this reason at least a candidate for criticism, it is not obviously a way of *wronging* the person praised. While the hypocritical blamer intuitively merits blame, the unfamiliar praiser typically elicits a weaker form of disapproval, one that is furthermore apt to be characterized by ambivalence, at least to the extent that we think the person praising *means well*.

The familiarity condition enjoins us to praise in ways that allow for co-valuation. I have presented only a sketch of the idea that co-valuation (or jointly acknowledged value resonance) is the function of praise. Still, this thesis stands to explain two independently plausible ideas: first, that there is something *good* about successful praise; second, that there is something *bad* about praise without standing. To the extent that these ideas are shared, we should at least consider welcoming the familiarity condition into our practical repertoire. This is so even if a different justification for the familiarity condition should ultimately be preferable. In any case—and I will be content if only the following has been established— it can no longer be taken for granted that norms of "standing to blame" have

no analogue in the arena of praise.[38] Symmetry is exhibited, not only in the significance we attribute to the features of those who *are* responsible, but also in the features of persons who might hold them so.

---

[38] It is a question for future work whether there are conditions of 'standing to praise' apart from the familiarity condition. While one cannot be "complicit" in another's praiseworthy conduct, we can nonetheless ask whether it is appropriate for A to praise B when A plays a substantive instrumental role in bringing it about that B merits praise. I suspect the answer will depend on the way in which one is instrumental. The piano teacher is intuitively in a position to praise her pupil, despite her role in the latter's success. But, extrapolating from what others have said about the blameworthiness of 'manipulated agents' (Pereboom 2001; Mele 2006), perhaps A will lack the standing to praise B for φ-ing if B performs φ as a result of A's covert control or manipulation of B, *even if* B is praise*worthy* from an appraiser-neutral standpoint, for φ-ing. See Russell (2010) and King (2015).

BIBLIOGRAPHY

Adams, Robert Merrihew (1985) "Involuntary Sins," *Philosophical Review* 94: 3-31.

Anderson, Elizabeth (1993) *Value in Ethics and Economics* (Harvard University Press).

Aristotle (2000) *Nicomachean Ethics* trans. R, Crip (Cambridge University Press).

Arpaly, Nomy and Schroeder, Timothy (2014), *In Praise of Desire* (Oxford University Press).

Arpaly, Nomy (2002) "Moral Worth" *Journal of Philosophy* 99: 223-245.

Arpaly, Nomy (2003) *Unprincipled Virtue: An Inquiry into Moral Agency* (Oxford University Press).

Arpaly, Nomy (2006) *Merit, Meaning, and Human Bondage: An Essay on Free Will* (Princeton University Press).

Bell, Macalester (2013) "The Standing to Blame" in *Blame: Its Nature and Norms*, eds. J. Coates and N. Tognazzini (Oxford University Press).

Bell, Macalester (2013), *Hard Feelings: The Moral Psychology of Contempt* (Oxford University Press).

Benn, Claire (2017) "Supererogation, optionality and cost" *Philosophical Studies*

Bennett, Christopher (2002) "The Varieties of Retributive Experience" *Philosophical Quarterly* 52(207):145-163.

Bennett, Jonathan (1980) "Accountability" in ed. Z. Van Straaten, *Philosophical Subjects: Essays Presented to P. F. Strawson* (Clarendon Press).

Bennett, Jonathan (2008 [1980]) "Accountability (II)" in eds. M. McKenna and P. Russell, *Free Will and the Reactive Attitudes: Perspectives on P.F. Strawson's "Freedom and Resentment"* (Ashgate).

Bentham, Jeremy (1781/1988), *The Principles of Morals and Legislation (*Prometheus Books).

Berger, Fred (1975) "Gratitude" *Ethics* 85: 298-309.

Björnsson, Gunnar and Karl Persson (2012) "The Explanatory Component of Moral Responsibility" *Noûs* 46.2: 326-354.

Bok, Hilary (1998) *Freedom and Responsibility* (Princeton University Press).

Bovens, Luc (1999) "The Value of Hope" *Philosophy & Phenomenological Research* 58: 667-81.

Bradford, Gwen (2015) *Achievement* (Oxford University Press).

Brownlee, Kimberley (2010) "Moral Aspirations and Ideals" *Utilitas* 22.3: 241-257.

Calhoun, Chesire (2004) "Common Decency" in ed. C. Calhoun, *Setting the Moral Compass* (Oxford University Press).

Callard, Agnes (2016) "Proleptic Reasons" *Oxford Studies in Metaethics* (vol. 11), ed. R. Shafer Landau (Oxford University Press).

Callard, Agnes (2018) "The Reason to be Angry Forever" in eds. M. Cherry and O Flanagan, *The Moral Psychology of Anger* (Rowman & Littlefield).

Carlsson, Andreas Brekke (2017) "Blameworthiness as Deserved Guilt" *Journal of Ethics* 21: 89-115.

Carr, David (2013) "Varieties of Gratitude" *Journal of Value Inquiry* 46: 17-28.

Carter, Phillip (2011) "For veterans, is 'thank you for your service' enough?" *The Washington Post,* Nov. 4— URL: https://www.washingtonpost.com/opinions/for-veterans-is-thank-you-for-your-service enough/2011/11/03/gIQA67hZmM_story.html?utm_term=.0bbcf57b810c

Chaucer, Geoffrey (1993) [1478] *The Canterbury Tales,* ed. R. Ecker and E. Crook (Hodge and Braddock).

Coates, Justin and Neal, Tognazzini (2013) "The Contours of Blame," in ed. Coates and Tognazzini, *Blame: Its Nature and Norms:* 3-26 (Oxford University Press).

Coates, Justin D. (2016) "The Epistemic Norm of Blame" *Ethical Theory and Moral Practice* 19.2: 457-473.

Cohen, G.A. (2006) "Casting the First Stone: Who Can, and Who Can't, Condemn the Terrorists?" *Royal Institute of Philosophy Supplement* 58: 113–136.

Clarke, Randolph (2013) "Some Theses on Desert" *Philosophical Explorations* 16(2): 153-164.

Clarke, Randolph (2016) "Moral Responsibility, Guilt and Retributivism" *The Journal of Ethics* 20: 121–137.

Dancy, Jonathan (2000) *Practical Reality* (Oxford University Press).

D'Arms Justin and Daniel Jacobson (2000) "The Moralistic Fallacy" *Philosophy and Phenomenological Research* 61: 66-90.

Darwall, Stephen (2006) *The Second-Person Standpoint: Morality, Respect, and Accountability* (Harvard University Press).

Darwall, Stephen (2007) "Moral Obligations and Accountability," *Oxford Studies in Metaethics* (vol. 2), ed. R. Shafer-Landau (Oxford University Press).

Darwall, Stephen (2012) "Bipolar Obligations" *Oxford Studies in Metaethics* (vol. 7), ed. R. Shafer-Landau (Oxford University Press).

Davidson, Donald (1980) [1978] "Hume's Cognitive Theory of Pride" in *Essays on Actions and Events* (Oxford: Clarendon Press).

Day, J.P. "Hope" (1969) *American Philosophical Quarterly* 6: 82-102.

Deonna, Julien and Fabrice Teroni (2012) *The Emotions: A Philosophical Introduction* (Routledge).

de Sousa, Ronald (1987) *The Rationality of Emotion*, Cambridge (MIT Press).

Downie, R.S. (1963) "Hope" *Philosophy and Phenomenological Research* 24: 248-251.

Driver, Julia (1992) "The Suberogatory" *Australasian Journal of Philosophy* 70(3): 286-295.

Ekman, Paul (2003) *Emotions Revealed* (Times Books).

Eshleman, Andrew (2014) "Worthy of Praise: Responsibility and Better-than-Minimally Decent Agency," in *Oxford Studies in Agency and Responsibility*, Vol. II: 'Freedom and Resentment' at 50, eds. D. Shoemaker and N. Tognazzini (Oxford University Press).

Eylon, Yuval and David Heyd (2008) "Flattery" *Philosophy and Phenomenological Research* 3: 685-704.

Feinberg, Joel (1966) "Duties, Rights, and Claims" *American Philosophical Quarterly* 3: 137–144.

Feinberg, Joel (1970) "Justice and Personal Desert" in *Doing and Deserving: Essays in the Theory of Responsibility* (Princeton University Press).

Feinberg, Joel (1986) *Harm to Others* (Oxford University Press).

Feldman, Fred and Brad Skow (2015) "Desert" ed. E. Zalta, *Stanford Encyclopedia of Philosophy*.

Fisher, John and Ravizza, Mark (1999) *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge University Press).

Fletcher, William A. (1988) "The Structure of Standing" *Yale Law Journal* 98: 221-291.

Foster Wallace, David (1996) *Infinite Jest* (Little, Brown and Company).

Frankfurt, Harry G. (1971) "Freedom of the Will and the Concept of a Person" *Journal of Philosophy* 68(1): 5-20.

Frankfurt, Harry G. (1982) "The importance of what we care about" *Synthese* 53: 252-272.

Franzén, Nils (2018) "Aesthetic Evaluation and First-Hand Experience" *Australasian Journal of Philosophy*, https://doi.org/10.1080/00048402.2018.1425729

Friedman, Marilyn (2013) "How to Blame People Responsibly" *The Journal of Value Inquiry* 47: 271-284.

Glover, Jonathan (1970) *Responsibility* (Routledge & Kegan Paul).

Grice, H. P. (1975) "Logic and conversation" in eds. J.L. Morgan & P. Cole, *Syntax and Semantics 3: Speech Acts*: 41–58 (Academic Press).

Gulliford, L. Morgan, B., and Kristjansson, K. (2013), "Recent Work on the Concept of Gratitude in Philosophy and Psychology," *Journal of Value Inquiry* 47: 285-317

Gulliford, L. and Morgan, B. (2014), "Measuring and Understanding Gratitude: A Theoretical and Empirical Approach" http://www.jubileecentre.ac.uk/485/papers/conference-papers/can-virtue-be-measured-papers/

Haji, Ishtiyaque (1997) "An Epistemic Dimension of Blameworthiness" *Philosophy and Phenomenological Research* 57: 523-44.

Hartman, Robert (2016) "Against luck-free moral responsibility" *Philosophical Studies*: 173: 2845–2865.

Helm, Bennett (2010) *Love, Friendship, and the Self: Intimacy, Identification, and the Social Nature of Persons* (Oxford University Press).

Helm, Bennett (2017) *Communities of Respect: Grounding Responsibility, Authority, and Dignity* (Oxford University Press).

Hemon, Aleksandar (2016) "Six Writers on the Genius of Marcel Proust" [Hustvedt, White, Aciman, Prose, Hemon, Mendelsohn] *Literary Hub*, http://lithub.com/six-writers-on-the-genius-of-marcel-proust/.

Heyd, David (1982) *Supererogation: Its Status in Ethical Theory* (Cambridge University Press).

Hieronymi, Pamela (2001) "Articulating an Uncompromising Forgiveness" *Philosophy and Phenomenological Research* 62(3): 529-555.

Hieronymi, Pamela (2004) "The Force and Fairness of Blame" *Philosophical Perspectives* 18: 115-148.

Higdon, Hal (1999) [1975] *Leopold and Loeb: The Crime of the Century* (University of Illinois Press).

Hoberman, J. (2012) "Action and Transaction: Hoberman on "the Kid With a Bike"" *Blouin Art Info*, March 15, http://www.blouinartinfo.com/news/story/763223/action-and-transaction-hoberman-on-%E2%80%9Cthe-kid-with-a-bike%E2%80%9D

Hohfeld, Wesley Newcomb (1923) *Fundamental Legal Conceptions*, ed. W. Wheeler Cook (Yale University Press).

Honderich, Ted (1988) *The Consequences of Determinism (*Oxford University Press).

Horgan, Terrance and Mark Timmons (2010) "Untying a knot from the inside out: Reflections on the 'paradox' of supererogation. *Social Philosophy and Policy Foundation* 27: 29-63.

Hume, David (1967) [1738] *A Treatise of Human Nature,* ed. L. A. Selby-Bigge (Oxford: Clarendon Press).

Hurley, Elisa and Macnamara, Coleen (2011) "Beyond Belief: Toward a Theory of Reactive Attitudes," *Philosophical Papers* 39: 373-399.

Isenberg, Arnold (1980) [1949] "Natural Pride and Natural Shame" in ed. A. Rorty, *Explaining Emotions* (University of California Press).

Kafka, Franz (1949), *The Diaries of Franz Kafka 1914-1923*, ed. M. Brod (Schocken).

Kane, Robert (1996) *The Significance of Free Will* (Oxford University Press).

Kant, Immanuel (1785/1998), *Groundwork of the Metaphysics of Morals*, trans. M. Gregor (Cambridge University Press).

Kant, Immanuel (1797/1998), *Metaphysics of Morals,* trans. M. Gregor (Cambridge University Press).

Kenner, Lionel (1967) "On Blaming" *Mind* 76.302: 238–249.

King, Matt (2014) "Two Faces of Desert" *Philosophical Studies* 169: 401-424.

King, Matt (2015) "Manipulation Arguments and the Moral Standing to Blame" *Journal of Ethics and Social Philosophy* 9: 1-20.

Kolodny, Niko (2003) "Love as a Valuing Relationship" *Philosophical Review* 112: 135-189.

Kristjánsson, Kristján (2002) *Justifying Emotions: Pride and Jealousy* (Routledge).

Landy, Joshua (2004) *Philosophy as Fiction: Self, Deception, and Knowledge in Proust* (Oxford University Press).

Latus, Andrew (2000) "Moral and Epistemic Luck" *Journal of Philosophical Research* 25: 149–172.

Lazarus, Richard (1991) *Emotion and Adaptation* (Oxford University Press).

Levy, Neil (2005) "The Good, The Bad and The Blameworthy" *Journal of Ethics & Social Philosophy* 1(2): 1-14.

Lewis, Michael (2000) "Self-Conscious Emotions: Embarrassment, Pride, Shame, and Guilt" in M. Lewis and J.M. Haviland-Jones eds. *Handbook of Emotions*, 2nd ed. (Gulliford Press).

Macnamara, Coleen (2011) "Holding others responsible" *Philosophical Studies* 152: 81-102.

Macnamara, Coleen (2013) ""Screw you!" & "thank you"" *Philosophical Studies* 165: 893-914.

Macnamara, Coleen (2015a) "Reactive Attitudes as Communicative Entities" *Philosophy and Phenomenological Research* 88: 546-569.

Macnamara, Coleen (2015b) "Blame, Communication, and Morally Responsible

Agency" in eds. R. Clarke, M. McKenna, and A. Smith, *The Nature of Moral Responsibility: New Essays* (Oxford University Press).

Manela, Tony (2016) "Gratitude and Appreciation" *American Philosophical Quarterly* 53: 281-294.

Markovits, Julia (2010) "Acting for the Right Reasons" *Philosophical Review* 119: 201-242.

Markovits, Julia (2012) "Saints, heroes, sages, and villains" *Philosophical Studies* 158: 289-311.

Martin, Adrienne (2013) *How We Hope: A Moral Psychology* (Princeton University Press).

Mason, Michell (2011) "Blame: Taking it Seriously" *Philosophy and Phenomenological Research* 83: 473-481.

McAleer, Sean (2012), ''Propositional Gratitude'' *American Philosophical Quarterly* 49: 55–66

McKenna, Michael (2012) *Conversation and Responsibility* (Oxford University Press).

Mele, Alfred R. (2006) *Free Will and Luck* (Oxford University Press).

Mellema, Gregory (1991) *Beyond the call of duty* (State University of New York Press).

Nagel, Thomas (1979), "Moral Luck," in *Mortal Questions* (Cambridge University Press).

Nelkin, Dana (2009) "Responsibility, Rational Abilities, and Two Kinds of Fairness Arguments" *Philosophical Explorations* 12(2): 151-165.

Nelkin, Dana (2013) "Desert, Fairness, and Resentment" *Philosophical Explorations* 16: 117-132.

Nelkin, Dana (2016) "Accountability and Desert" *Journal of Ethics* 20: 173-189.

Nichols, Shaun (2007), "After Incompatibilism: A Naturalistic Defense of the Reactive Attitudes" *Philosophical Perspectives*, 21: 405-428.

Ninan, Dilip (2014) "Taste Predicates and the Acquaintance Inference" *Semantics and Linguistic Theory* 24: 290–309.

Nussbaum, Martha (1994) *The Therapy of Desire (*Princeton University Press).

Nussbaum, Martha (2001) *Upheavals of Thought: The Intelligence of Emotions*

(Cambridge University Press).

Nussbaum, Martha (2016), *Anger and Forgiveness: Resentment, Generosity, Justice* (Oxford University Press).

Pereboom, Derk (1995), "Determinism *al Dente*", *Nous,* Vol. 29, No. 1 (1995), 21-45.

Pereboom, Derk (2001) *Living Without Free Will* (Cambridge University Press).

Pereboom, Derk (2007) "Hard Incompatibilism" in *Four Views on Free Will* (Fischer, Kane, Pereboom, Vargas) (Blackwell Press).

Pereboom, Derk (2013) "Free Will Skepticism, Blame, and Obligation" in eds. D.J. Coates and N.A. Tognazzini, *Blame: Its Nature and Norms*: 189-206 (Oxford University Press).

Pereboom, Derk (2014) *Free Will, Agency, and Meaning in Life* (Oxford University Press).

Pereboom, Derk (2017) "Responsibility, Regret, and Protest" in ed. D. Shoemaker, *Oxford Studies in Agency and Responsibility (Volume* 4) (Oxford University Press).

Pickard, Hannah (2014) "Irrational Blame" *Analysis* 73: 613-626.

Prinz, Jesse (2004) *Gut Reactions: A Perceptual Theory of Emotion* (Oxford University Press).

Proust, Marcel (1922) [1913] *Remembrance of Things Past: Swann's Way*, trans. C.K. Scott Moncrieff (Henry Holt and Company).

Radzik, Linda (2011) "On Minding Your Own Business: Differentiating Accountability Relations within the Moral Community" *Social Theory and Practice* 37: 574–598.

Richel, Matt (2015) "Please Don't Thank Me for My Service" *New York Times* Feb. 21— url: http://www.nytimes.com/2015/02/22/sunday-review/please-dont-thank-me-for-my-service.html?_r=0

Roberts, Robert (2003) *Emotion: An Essay in Aid of Moral Psychology* (Cambridge University Press).

Roberts, Robert (2009) "The Vice of Pride" *Faith and Philosophy* 26(2): 119-133.

Roberts, Robert (2013) *Emotions in The Moral Life* (Cambridge University Press).

Roberts, Robert (2015) "The Normative and Empirical in the Study of Gratitude" *Res*

Philosophica 92(4): 883-914.

Rosen, Gideon (2004) "Skepticism about Moral Responsibility" *Philosophical Perspectives* 18:295-313.

Rosen, Gideon (2015) "The Alethic Conception of Moral Responsibility" in eds. R. Clarke, M. McKenna, and A. Smith (Oxford University Press

Russell, Paul (1992) "Strawson's Way of Naturalizing Responsibility" *Ethics* 10(2): 287-302.

Russell, Paul (1995) *Freedom and Moral Sentiment: Hume's Way of Naturalizing Responsibility* (Oxford University Press).

Russell, Paul (1999) "Smith on Moral Sentiment and Moral Luck" *History of Philosophy Quarterly* 16: 37-58.

Russell, Paul (2004) "Responsibility and the Condition of Moral Sense" *Philosophical Topics*, 32: 287-305.

Russell, Paul (2008) "Free Will, Art and Morality" *The Journal of Ethics* 12(3): 307-325.

Russell, Paul (2010) "Selective Hard Compatibilism" in J. Campbell, M. O'Rourke, and H. Silverstein eds., *Action, Ethics, and Responsibility*: 149-173 (MIT Press).

Russell, Paul (2013) "Responsibility, Naturalism, and "The Morality System"" in *Oxford Studies in Agency and Responsibility (Volume 1)* ed., D. Shoemaker (Oxford University Press).

Russell, Paul (2017a) "Free Will Pessimism" in P. Russell, *The Limits of Free Will* (Oxford University Press).

Russell, Paul (2017b) "'Can we Talk?' [Review of Michael McKenna, *Conversation and Responsibility*" *Philosophical Review* 126: 285-95.

Scanlon, T. M. (1998) *What we owe to each other* (Harvard University Press).

Scanlon, T.M. (2008) *Moral Dimensions: Permissibility, Meaning, Blame* (Belknap Press).

Scanlon, T.M. (2013) "Interpreting Blame" in eds. J. Coates and N. Tognazzini, *Blame: Its Nature and Norms* (Oxford University Press).

Scheffler, Samuel (1997) "Relationships and Responsibilities" *Philosophy & Public Affairs* 26: 189-209.

Scheffler, Samuel (2010) "Valuing" in his *Equality and Tradition: Questions of Value in Moral and Political Theory*: 15-40 (Oxford University Press).

Schlick, Moritz (1939) *Problems of Ethics*, trans. D. Rynin (Prentice-Hall).

Searle, John (1975) "Indirect Speech Acts", in eds. J.L. Morgan & P. Cole, *Syntax and Semantics 3: Speech Acts*: 59–82 (Academic Press).

Searle, John (1979) *Expression and Meaning: Studies in the Theory of Speech Acts* (Cambridge University Press).

Searle, John R. and Daniel Vanderveken (1985) *Foundations of Illocutionary Logic* (Cambridge University Press).

Sher, George (1987) *Desert* (Princeton University Press).

Sher, George (2001) "Blame for Traits" *Noûs* 35.1: 146-161.

Sher, George (2006) *In Praise of Blame* (Oxford University Press).

Sherman, Nancy (2010) *The Untold War: Inside the Hearts, Minds, and Souls of Our Soldiers* (W.W. Norton & Company).

Sherman, Nancy (2015) *Afterwar: Healing the Moral Wounds of Our Soldiers* (Oxford University Press).

Seneca, Lucius Annaeus (56-64 CE/2011), *On Benefits*, trans. M. Griffin and B. Inwood (University of Chicago Press).

Shoemaker, David (2011) "Attributability, Answerability, and Accountability: Toward A Wider Theory of Moral Responsibility" *Ethics* 121: 602-632.

Shoemaker, David (2013) "Qualities of Will," *Social Philosophy & Policy* 30: 95-120.

Shoemaker, David (2015) "Ecumenical Attributability" in eds., R.Clarke, M. McKenna, A. Smith, *The Nature of Responsibility* (Oxford University Press).

Shoemaker, David (2015) *Responsibility from the Margins* (Oxford University Press).

Shoemaker, Sydney (1968) "Self-Reference and Self-Awareness," *Journal of Philosophy* 65: 555-567.

Sidgwick, Henry (1981 [1907]) *The Methods of Ethics*, 7th ed. (Hackett).

Smilansky, Saul (1997) "Should I Be Grateful to You for Not Harming Me?" *Philosophy and Phenomenological Research* 57: 585-97.

Smith, Angela (2005) "Responsibility for Attitudes: Activity and Passivity in Mental Life" *Ethics* 115.2: 236-271.

Smith, Angela (2007) "On Being Responsible and Holding Responsible" *The Journal of Ethics* 11: 465- 484.

Smith, Angela (2008) "Control, Responsibility, and Moral Assessment" *Philosophical Studies* 138: 367-392.

Smith, Angela (2011) "Guilty Thoughts," in Carla Bagnoli (ed.), *Morality and the Emotions* (Oxford University Press).

Smith, Angela (2013) "Moral Blame and Moral Protest" in eds. J. Coates and N. Tognazzini, *Blame*: *Its Nature and Norms* (Oxford University Press).

Solomon, Robert C. (1976) *The Passions* (Anchor Press/Doubleday).

Strawson, Peter (1962/1974), "Freedom and Resentment," in *Freedom and Resentment and other Essays (*Methuen & Co Ltd).

Strawson, Peter Frederick (1970 [1961]) "Social Morality and Individual Ideal" in eds. G. Wallace & A.D.M. Walker, *The Definition of Morality* (Methuen & Co Ltd.).

Strohminger, Nina and Shaun Nichols (2014) "The Essential Moral Self" Cognition 131: 159-171.

Talbert, Matthew (2012) "Moral Competence, Moral Blame, and Protest" *Journal of Ethics* 16: 89-109.

Taylor, Gabriele (1985) *Pride, Shame, and Guilt: Emotions of Self-Assessment* (Clarendon Press).

Tosi, Justin and Brandon Warmke (2016) "Moral Grandstanding" *Philosophy & Public Affairs* 44: 197-217.

Tracy, Jessica and Richard Robins (2004) "Putting the self into self-conscious emotions: A theoretical model" *Psychological Inquiry 15:* 103–125.

Tracy, Jessica and Richard Robins (2007) "The Psychological Structure of Pride: A Tale of Two Facets" *Journal of Personality and Social Psychology* 92: 506-525.

Urmson, James Opie (1958) "Saints and Heroes", in ed. A. Melden, *Essays in Moral Philosophy* (University of Washington Press).

Vargas, Manuel (2007) "Revisionism" in *Four Views on Free Will* (Fischer, Kane,

Pereboom, Vargas) (Blackwell Press).

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

Wallace, R. Jay (2007) "Reasons, Relations, and Commands" *Ethics* 118: 24-36.

Wallace, R. Jay (2010) "Hypocrisy, Moral Address, and the Equal Standing of Persons" *Philosophy and Public Affairs* 38: 307-341.

Wallace, R. Jay (2012), "Justification, Regret, and Moral Complaint: Looking Forward and Looking Backward on (and in) Human Life," *Luck, Value, and Commitment: Themes from the Ethics of Bernard Williams,* eds. U. Heuer and G. Lang (Oxford University Press).

Wallace, R. Jay (2013) *The View from Here: On Affirmation, Attachment, and the Limits of Regret* (Oxford University Press).

Wallace, Jay (2014) "Reasons, Policies, and the Real Self" in eds. M. Vargas and G. Yaffe, *Rational and Social Agency: Philosophy of Michael Bratman* (Oxford University Press).

Wallace, R. Jay (ms.) *The Moral Nexus* [draft of March 28, 2017]

Walker, A.D.M. (1980) "Gratitude and Gratefulness" *Proceedings of the Aristotelian Society* 81:39–55.

Walker, A.D.M. (1988) "Political Obligation and the Argument from Gratitude" *Philosophy & Public Affairs* 17:3, 191-211.

Watson, Gary (1975) [2004] "Free Agency" in G. Watson, *Agency and Answerability: Selected Essays* (Clarendon).

Watson, Gary (1987) [2004] "Responsibility and the Limits of Evil" in G. Watson, *Agency and Answerability: Selected Essays* (Clarendon).

Watson, Gary (1996) [2004] "Two Faces of Responsibility" in G. Watson, *Agency and Answerability: Selected Essays* (Clarendon).

Watson, Gary (2011) "The Trouble with Psychopaths" in eds. R. J. Wallace, R. Kumar, S. Freeman, *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon* (Oxford University Press).

Wedgwood, Ralph (2009) "The "Good" and the "Right" Revisited" *Philosophical Perspectives* 23: 499-519.

Wertheimer, Roger (1998) "Constraining Condemning" *Ethics* 108: 489–501.

Wheatley, J. (1958) "Wishing, and Hoping" *Analysis* 18: 121-131.

Williams, Bernard (1981) "Moral Luck," in *Moral Luck* (Cambridge University Press).

Wolf, Susan (1981) "The Importance of Free Will", *Mind* Vol. XC, 386-405.

Wolf, Susan (1987) "Sanity and the Metaphysics of Responsibility" in ed. F. Schoeman, *Responsibility, Character, and the Emotions* (Cambridge University Press).

Wolf, Susan (1990) *Freedom within Reason* (Oxford University Press).

Wolf, Susan (2001) "The Moral of Moral Luck," *Philosophic Exchange* 31: 4–19.

Wolf, Susan (2015) "Responsibility, Moral and Otherwise" *Inquiry* 58: 127-142.

Wood, A., Maltby, J., Stewart, N., and Joseph, S. (2008) ''Conceptualizing Gratitude and Appreciation as a Unitary Personality Trait'' *Personality and Individual Differences* 44: 621-63

Zimmerman, Michael J. (1988) *An Essay on Moral Responsibility* (Rowman & Littlefield).

Zimmerman, Michael J. (1996) *The Concept of Moral Obligation* (Cambridge University Press).

Zimmerman, Michael J. (2002) "Taking Luck Seriously," *The Journal of Philosophy*, 99: 553–576.

Zimmerman, Michael J. (2015) "Varieties of Moral Responsibility" in eds. R. Clarke, M. McKenna, and A.M. Smith, *The Nature of Moral Responsibility: New Essays* (Oxford University Press).