Moralized theories of friendship have fallen out of favor. Philosophers have set Aristotle aside to examine goods of friendship that are achievable in the absence of goodness of character. We are invited to attend to the ways in which friendship requires openness to being shaped by another, while backing away from concerns about how our moral characters may be shaped through the relationships we choose to form and deepen.¹ This turn is good in many ways for the philosophy of friendship, but I have come to think that it is unfortunate for moral theory insofar as it has dampened discussion of moral friendship. While I do not doubt that there is a sense in which a morally corrupt or immature person can be a good friend, such a person is certainly not a candidate for a moral friendship, understood as a friendship between people who share a sense of and commitment to moral principles. Moral theory need not be embarrassed of its continued interest in the category of moral friendship.

My aim in this paper is to develop an account of moral friendship that clarifies the role of this form of relationship in the pursuit of moral self-perfection. The view I offer might be considered a reformatory reading of Kant’s presentation of moral friendship in the Doctrine of Virtue, one that reveals what Kant’s view could have been if Kant did not have certain limitations in his thinking about relationality. I seek to build upon the most insightful aspects of Kant’s account, in particular its emphasis on mutual trust, while also correcting for its problems and deficiencies. Kantian ethics is a particularly promising framework within which to consider moral friendship because of Kant’s prioritization of the duty of moral self-perfection, attention to both individual and social dimensions of moral striving, and acute sense of the social obstacles to this striving. My account will explain how trusting and being trusted in the way that Kant describes supports an agent’s moral self-cultivation beyond increased self-knowledge and refinement of judgment. I attend not only to the ideal of complete confidence between two people but also to the development of moral friendship,

¹. See, for example, Cocking and Kennett (1998) and (2000) and Nehamas (2016).
which requires the cultivation of trust and trustworthiness on the part of both parties.

A full Kantian account of moral friendship, I will argue, needs to address its two main roles with respect to the end of moral self-perfection: reducing obstacles to moral improvement and creating positive possibilities for moral development. In Section I, I offer an overview of Kant’s conception of moral friendship that clarifies the ways in which the cultivation of such a friendship removes obstacles to virtue by combatting the unsociable side of our unsocial sociability. In Section II I shift from reconstructing Kant’s views to constructing my own Kantian theory of moral friendship. I detail specific modes of pursuing virtue made possible by a relationship of robust reciprocal trust, many of which are not recognized by Kant because of his preoccupation with prudential concerns and his limited attention to relational dimensions of self-cultivation. Taken together, my arguments reveal the central role that trust relations play in rendering progress towards moral self-perfection possible within an ethical theory that acknowledges social antagonism and fragmentation as among the most hazardous threats to individual moral development.

I. Friendship as Refuge: Distrust and the Social Passions

Kant’s account of moral friendship in the Doctrine of Virtue of the Metaphysics of Morals follows his presentation of friendship as an ideal of equal, reciprocal love and respect that we have a duty to pursue. The defining feature of moral friendship is mutual trust which allows two people to reveal their thoughts and feelings freely, without fear of judgment or loss of respect. Kant understands communication of one’s thoughts as a need, frustration of which leaves a person constrained, isolated, and “shut up in himself” (MS, 6:472).

The fact that a person must already be “upright,” intelligent and trustworthy to be a suitable participant in moral friendship may suggest a somewhat limited role for friendship in the pursuit of self-perfection (C, 27:429). One must already be “good enough” to benefit from a trusting relationship with another good enough person. But

we might think of preparing oneself for friendship as part of the fulfillment of the duty of friendship, such that this duty overlaps with the duty of moral self-perfection. “Good enough” people are still imperfect and thus moral friendship is not an ultimate achievement, but rather a way of being in relation that creates the conditions for particularly effective modes of moral striving. Unlike Aristotle, Kant asks us to imagine the parties to moral friendship as people in a transitional state who, while perhaps having significant room for improvement, are morally capable of deserving another’s trust.

We make ourselves worthy of trust, Kant tells us, by avoiding malicious and deceptive action and by being upright and candid (C, 27:429). Here we see that to some extent cultivating the traits that make us worthy of friendship requires us to engage in the characteristic activity of moral friendship, the sharing of our thoughts. Kant argues that the mutual enjoyment of sharing thoughts “is the foundation of open-heartedness, animus apertus sinceritas aperta,” that is, a candid and sincere disposition (V, 27:678). In turn, openheartedness allows “the so needful sharing of feelings and thoughts, the necessary enlargement of our various perfections, and the closer bonding with the friend” (V, 27:679).

Moral friendship definitely makes life more pleasant and it helps us improve upon our existing abilities (enlarging our perfections), yet we may wonder whether moral friendship provides “the icing on the cake” of established goodness or has a more significant role in moral striving. In the Collins lectures, Kant remarks, “By friendship we cultivate virtue in little things” (C, 27:430). Are openheartedness, willingness to trust, and trustworthiness just minor virtues?2 We might think that

2. Some philosophers argue that trustworthiness is not a virtue at all because whether trustworthiness serves virtue or vice depends on what the person is trusted to do. Further, trust is good or bad depending on how wisely it is placed. I will not take a position on the general question of whether morality constrains trustworthiness internally or externally, as my focus is on trust and trustworthiness within the context of moral friendship. In this context trustworthiness requires being worthy of participation in the friend’s moral development; it is impossible to be worthy of such participation if one does not respect moral law.
there is a good reason why Kant would not want to make friendship too central to the pursuit of virtue, namely that it is not entirely within the individual’s control. If no one is trustworthy, it is advisable to remain reserved. The question, however, is whether constant concealment of one’s thoughts and feelings is a big or a small cost, morally speaking.

The Vigilantius lectures represent mutual trust as an achievement that takes place against a background of general distrust and reserve that easily devolves into envious rivalry. Man has not just a propensity, “but a well-considered inclination, to hide from his fellow-man in sharing his inner state,” for prudence dictates that we mask our flaws so as to avoid losing another’s respect. Kant asserts his doctrine of unsocial sociability, identifying social antagonism as a spur to the development of human powers, yet also a barrier to moral progress. Our competitiveness, rooted in anxiety about our own status vis-à-vis others, may even lead us to resent the other’s virtues and exaggerate his faults. The “greatest obstacle to friendship” is envy of other people’s merits (both moral and non-moral), but clearly envy is also among the greatest obstacles to moral self-perfection, the pursuit of which requires comparing oneself to moral law as opposed to other people. Kant further laments that “the suspicion aroused by emulation ensures that the other never discloses himself completely; the evil is that we virtually never regard the other as a friend, but rather as an opponent.” We cannot trust because we are not willing to give the other person power to harm or humiliate us.

Our concern for prudence does not just deter us from cultivating trust and openheartedness; it pits us against other people. Distrust goes beyond agnosticism about whether to trust because it constitutes a negative judgment about the other’s intentions or competence.

Indeed, at least at times, Kant takes the view that most people have a “base cast of mind” and many are “indiscreet or incapable of judging and distinguishing what may or may not be repeated” (MS, 6:472). He tells us that as a general rule we should regard the friend as a potential future enemy, trusting only with caution. Yet to trust with this level of restraint actually amounts to withholding trust, since to trust is to “forgo precaution” in our dealings with the other person. You cannot lower your guard while keeping it up. The idea that we should withhold both trust and distrust when we lack evidence of the other’s uprightness and good judgment sounds sensible enough, but it is worth considering how an inability to fully or robustly trust others impacts our moral striving. Only a person who has “meta-trust,” understood as “trust in trust-involving relationships,” is in a position to cultivate moral friendship with another or cultivate a more inclusive moral community. As Annette Baier notes, trust and mistrust “tend to be self-fulfilling, and tend to be contagious.”

Kant believes that the best friendships are free from the “emulation” he so vividly highlights. These relationships are rare, however, because even “well-meaning people” have “a drive constantly to perfect oneself in comparison with others” (V, 27:680). Kant makes a similar claim in Part III of Religion Within the Boundaries of Mere Reason where he introduces the notion of the Ethical State of Nature. Reading this passage and his reflections on friendship together is telling. According to Kant, the lure of prioritizing self-love over the moral law will hound the striving individual “however much he exerts himself,” and the only recourse available is to leave the Ethical State of Nature and form an Ethical Commonwealth, that is, a unified moral community (R, 6: 93). No degree of strength achieved through the individual’s own performance can make one trustworthy. Kant is proud of his “manhood,” and he is aware that he is not the only one who is proud of his “manhood.” Kant’s pride is not the pride of the man who is proud of his own manhood, but the pride of the man who is proud of his humanity. Kant knows that he is not the only one who is proud of his humanity, but he is aware that he is not the only one who is proud of his humanity. Kant is proud of his humanity, and he is aware that he is not the only one who is proud of his humanity.


3. In §36 of the Doctrine of Virtue Kant groups envy together with ingratitude and malice under the category of “vices of hatred for human beings.” Kant writes, “In these vices, however, hatred is not open and violent but secret and veiled, adding meanness to one’s neglect of duty to one’s neighbor, so that one also violates a duty to oneself” (MS, 6:458).


8. Kant refers to this as “the principle of evil” in this passage and elsewhere it is referred to as the “propensity to evil.”

9. The Ethical Commonwealth refers to a society with robust, non-coercive
moral striving can make her immune to the social passions, which are aroused by the proximity of other people.

Envy, addiction to power, avarice, and the malignant inclinations associated with these, assail his nature, which on his own is undemanding, as soon as he is among human beings. Nor is it necessary to assume that these are sunk into evil and are examples that lead him astray: it suffices that they are there, that they surround him, and that they are human beings, and they will mutually corrupt each other's moral dispositions and make one another evil (R, 6:94).

This passage serves as part of an argument for the claim that effective moral striving necessarily has a social dimension; one must “strive with all his might to extricate himself” from the Ethical State of Nature because this state of social division intensifies the unsociable side of our unsocial sociability. There is a direct parallel between the Ethical State of Nature and the Juridical State of Nature, for “In these two [states of nature] each individual prescribes the law to himself, and there is no external law to which he, along with the others, recognizes himself to be subject” (R, 6:95). In the case of the Ethical State of Nature, the relevant laws are the laws of virtue and the absence of an external law means that there are no institutions that embody and promote an interpretation of the requirements of these laws. When a group of people is not united by a shared understanding of the determinate content of duties of virtue there is “public feuding between the principles of virtue and a state of inner immorality,” which encourages individuals to subordinate morality to self-love (R, 6:97).10

Kant stresses that in order for two people to develop a moral friendship they must share principles “of understanding and morality” (C, 27:429); friends’ judgments need not be the same, but the conditions for harmony of judgment must be in place. Parties to the best moral friendships trust each other, do not compete, and share moral principles. Although Kant never mentions friendship in Part III of the Religion, the concern he expresses here regarding the lack of shared principles raises questions about how his view of social fragmentation fits together with his conception of moral friendship. Does the fact that moral friends share principles mean that they have left the Ethical State of Nature? To what extent does moral friendship solve the problem of the Ethical State of Nature, which is essentially a matter of combating the manifestations of our unsociability that undo virtue?

Answering these questions will require a fuller interpretation of Kant’s explanation of the Ethical State of Nature and the process of creating a unified ethical community. We need to say why people in an Ethical State of Nature do not share principles and why this so severely hampers individual moral striving. As with the Juridical State of Nature, indeterminacy is a fundamental problem in the Ethical State of Nature. Moral law fixes duties of virtue at a general level, but many uncertainties remain regarding things like how to balance the pursuit of required ends with that of permissible non-moral ends and the criteria for the fulfillment of the duty of moral self-perfection.11 The absence of shared standards for virtue muddies moral self-assessment and action for positive ends, which leaves the individual disoriented with respect to the end of moral self-perfection.

In the Ethical State of Nature an individual is not deprived of the moral law, but she is deprived of the social conditions conducive to both bringing the passions under reason’s control and to the establishment of firm commitment to the ongoing pursuit of positive moral ends. Moral friendship provides a needed resource in that the assurance of one’s own worth provided by the friend alleviates the craving

---

10. Anderson-Gold’s (1986) gloss on this line is that “the divisions between self and other are reflected within the individual as divisions within his or her will” (p. 30).

to mollify uncertainty about one’s own moral worth through regression to comparative social status as an evaluative standard. The idea here is that when we do not share principles with others, when “each individual is his own judge” as regards his own moral character, we have no way to confirm our self-assessments. The friend does not establish my self-assessments as objectively valid, but she does buttress my sense of worth.

The fact that the friend trusts me with her thoughts shows very directly that she considers me good: I am regarded as worthy of trust, meaning I am perceived as upright, of good judgment, etc. To be trusted is to be respected; in trusting me, the friend communicates a positive evaluation of my moral worth and my judgment. My thoughts matter to the friend, who takes them into consideration, even revising some of her own judgments in light of my insights. In this way, moral friendship gives us respite from rivalry and from competitive self-assessment based on social standing. Friendship is one’s “refuge in this world from distrust of his fellows” (C, 27:428).

We have seen in both contexts — in discussions of friendship and of the Ethical State of Nature — Kant stresses that vulnerability to envy is shared even by those who mean well. Moral friendship creates space for moral concern in that it allows the friends to step back from assessing themselves in competitive terms. Kant highlights our hypersensitivity to social status, but acknowledges that successful friendship eases our social anxieties.

Here again, I would stress that moral friendship is not an ultimate achievement. The overcoming of competitiveness is partially constitutive of trust, but also facilitated by trust. More specifically, the kind and level of competitiveness that would make one untrustworthy must be overcome in order to develop a moral friendship, but reciprocal trust does not require that one have totally overcome the vices of comparison. Moral friends are “good enough” but not necessarily excellent and certainly not perfect. Trust is what allows us to come to know how the vices of comparison and the assumptions they fuel hurt our friends, and love is what motivates us to overcome these assumptions. Trust is also what allows the reality of the other and her struggles to become known, which tends to undermine forms of competition that depend upon illusions. As a moral friendship develops, relative social status becomes less salient than status vis-à-vis the friend. In such a relationship, the mere presence of the other person does not provoke the passions because the background of competition has been dismantled in ways made possible by trust. Thus, reciprocal trust shelters the individual’s predisposition to goodness, which comes under attack by the passions in social relations wherein distrust is the default.

Kant considers the morally well-disposed, yet imperfect individual in isolation immediately before the previously cited passage describing his attack by “malignant inclinations” upon encountering others:

> His needs are but limited, and his state of mind in providing for them moderate and tranquil. He is poor (or considers himself so) only to the extent that he is anxious that other human beings will consider him poor and will despise him for it (R, 6:94).

The individual’s tranquility is disrupted by the scrutiny and judgments of distrusted others. It is largely lack of trust that causes us to corrupt each other’s moral dispositions through the arousal of the social passions. Kant does not represent friendship as mutually corrupting, for our trusted friends see us as equals deserving respect and love. As Annette Baier observes, “Mistrust can bring out the worst in the mistrusted, as trust can bring out the best.”13 In Part I of the Religion Kant speaks of the predisposition to humanity as a self-love that begins by seeking equal worth in the eyes of others, but develops into an “unjust desire to acquire superiority for oneself over others” because of fear that “that others might be striving for ascendancy” (R, 6:27). When we actually experience ourselves as equal to another, the desire for

12. For discussion of this point see D’Cruz (2015).
superiority dissolves.\textsuperscript{14} I do not need to best my best friend. Challenging Kant’s insistence on exact equality in friendship, Oliver Sensen asks, “Could it be that friendship lives off hope and trust as well as moments in which one experiences a unity with the other, rather than certainty about a permanent state?”\textsuperscript{15} I think the answer is yes. Kant recognizes the importance of trust to the free sharing of thoughts, but trust is also integral to the experience of oneself as a respected equal. In this way trust keeps the social passions at bay.

Here one may suspect that the suggested link between friendship and the calming of the social passions relies on Kant’s questionable idealization of moral friendship. Interestingly, Kant is open to accusations of being both too pessimistic and too optimistic and moralizing about friendship.\textsuperscript{16} Alexander Nehamas writes, “Even close friendships are often much less edifying than the rhetoric that surrounds the institution would have them be.”\textsuperscript{17} His criticism of the view that friendship is free from jealousy would certainly apply to my elaboration of Kant’s view. I may have overstated the freedom from the social passions associated with moral friendship as feelings of envy may arise, even where bonds of trust are robust. My conjecture is that where such bonds hold, intermittent feelings of envy will be benign as opposed to malicious.\textsuperscript{18}

Sharing sympathetically in the friend’s wellbeing is compatible with benign envy, understood as the wish to have what the friend has.

Consider Elena, the protagonist of Elena Ferrante’s My Brilliant Friend (2012), who envies her best friend Lila’s intellect, creativity, beauty and, eventually, wealth, yet consistently acts for her benefit. At one point, Elena is tempted to steer Lila toward breaking off her engagement to a successful man who has lifted her out of poverty.

Inside, it was what I truly wanted: to bring her back to pale, ponytailed Lila, with the narrowed eyes of a bird of prey, in her tattered dress. No more of those airs, that acting like the Jacqueline Kennedy of the neighborhood.

But, unfortunately for her and for me, it seemed a small-minded act. Thinking it would be for her good, I would not restore her to the bleakness of the Cerullo house ....\textsuperscript{19}

The common compliment “You make me want to be better” expresses painless benign envy and underscores its role in striving. In Elena’s case, benign envy fuels her academic efforts, which are rewarded when a teacher reads one of her essays aloud for an exam committee. Elena remarks, “And only as I listened did I realize what I had tried to do in those months whenever I had to write: to free myself from my artificial tones ... to try for a fluid and engaging style like Lila’s.” She adds, “Naturally it wasn’t Lila’s way of writing, it was mine.”\textsuperscript{20} Envy need not unseat beneficence and it can spur not mere imitation, but true individuality. Though this example is arguably one of non-moral striving, setting aside the notion of natural self-perfection as a duty to self, envy in the form of aspirational admiration may just as well motivate increased generosity or honesty as elegant writing. Even where painful feelings of envy arise within a moral friendship, robust mutual

\textsuperscript{14} Kolodny (2010) refers to this shift from the desire for equality to the desire for superiority as “egalitarian defensiveness” (p. 177). He notes that seeking superiority as a means to equality “is strictly speaking incoherent,” but can be made sense of given that being evaluated as morally superior is a particularly effective way to avoid being evaluated as morally inferior (p. 178). Kolodny considers Kant’s explanation of the predisposition to humanity inadequate as an explanation of inflamed amour-propre partly because Kant does not explain the basis for the belief that others seek to be evaluated as morally superior to oneself. The inability to trust other people is surely part of that explanation. Adding it in shows how Kant’s analysis can contribute to solving the puzzle of a moral psychology that explains both people’s actual moral weakness under current social conditions and potential moral strength under different social conditions.

\textsuperscript{15} Sensen (2013), p. 158.

\textsuperscript{16} Veltman (2004), for example, argues that Kant’s account of “unsocial sociability” damages his account of friendship (p. 226).

\textsuperscript{17} Nehamas (2016), p. 199.

\textsuperscript{18} For a succinct explanation see Smith (2015).

\textsuperscript{19} Elena Ferrante (2012), p. 310.

\textsuperscript{20} Ibid., p. 276.
trust and love limited by respect can reduce the duration and destructiveness of these feelings. A second question is whether the experience of moral friendship can meaningfully increase one's ability to govern the passions if the friends still distrust everyone else. Indeed, Kant maintains that the problems of the Ethical State of Nature are reproduced between societies and he also notes the dangers of tribalism in his lectures on friendship. Our principles battle their principles; we are better than them. Participation in one relationship that is free from competitive self-assessment may do little to combat envious rivalry in one's other relations. Close friendships may even facilitate the development of superiority complexes; we are better than everyone else.

The problem of tribalism is alleviated in the context of specifically moral friendship, as these friendships are less prone to morally problematic exclusion than friendships that are not well-constituted by commitment to moral principle. That said, moral friendship clearly cannot solve the problems associated with the Ethical State of Nature — if it did, we would not need to work toward the establishment of a unified ethical community. What moral friendship can do is temper the threat that the social passions pose to the striving agent's moral disposition and offer a resource that is available when an ethical commonwealth is not. Yes, friends must be careful not to "close their hearts" to others as they open them to each other, but insularity is surely not inevitable (C, 27:428). A relationship of reciprocal trust can help individuals govern their passions, mitigating the moral strain of living in the absence of a unified moral community.

Unlike affects, the passions alter an agent's deliberation about ends, undermining his ability to fulfill and deepen his moral commitments. "By corrupting or distorting the agent's practical thinking, the passions do damage to his basic predisposition to good, while the damage done by the affects is relatively impermanent and easier to control." Kant speaks of governing the passions and although he leaves this process somewhat mysterious, he indicates that it involves a variety of cognitive strategies in addition to the ability to anticipate the ways in which the passions may impact one's judgments. Attention to moral friendship makes the possibility of rational self-governance in the face of social fragmentation and antagonism less mysterious, revealing trust as a condition of proper moral self-understanding. The experience of communion and relative freedom from envy and competition within a moral friendship gives friends a reference point, an experience of self in social space that is not fraught. Relative freedom from the social passions in some interactions positions friends to develop strategies for reducing the influence of the passions in the context of less secure relationships.

II. Friendship as Communion: Sincerity, Being Known and Self-Governance

In this section I further establish the connection between moral friendship and self-governance in non-ideal social conditions, though my method will shift from the reconstruction (and moderate extension) of Kant's views to the construction of my own Kantian account of moral friendship. In Section I, I demonstrated that the basis for an explanation of the role of moral friendship in combatting the unsociable side of our unsocial sociability is there to be found in Kant's thought. This explanation is compelling and it is half of the story I want to tell about the role of robust reciprocal trust in moral striving. Yet because moral striving is a matter of both avoiding vice and pursuing virtue, a full


22. As Kant explains, "each particular society is only a representation or schema" of the ethical community, "for each of these societies can in turn be represented in relation to others of this kind, as situated in the natural state, with all the imperfections of the latter ..." (R, 6:96).

23. For discussion of this point see Nehamas (2017), p. 61.


Kantian account of moral friendship must address the positive contribution it makes to moral self-perfection. Kant sets us off in the right direction in that he makes communion central to his conception of moral friendship. An explanation of the role of communion in positive moral striving is the essential second half of my story, the half that requires going beyond Kant’s own thought.

What is communion and why does it matter for moral self-perfection? Communion is not a mystical state, but a form of relationality that depends upon open and successful communication. Communion does not just remove obstacles to moral improvement, such as feelings of malicious envy; rather, it is characterized by relations of sincerity and being-known that together underwrite self-governance. (1) Sincere expression and (2) the experience of being known within moral friendship alter one’s relation to oneself, thus changing not only the resources one has, but the position from which one strives for moral improvement. Further, (1) and (2) underwrite (3) self-governance because together they constitute the eschewal of deception of self and other, and a method of holding oneself responsible for the ongoing pursuit of the positive ends of reason.

(1) Sincerity

Let us begin, then, with an analysis of sincerity. I identify the connections between sincerity and the positive expression of commitment to three specific duties of virtue, namely, (i) truthfulness, (ii) respect and (iii) self-knowledge. In each case I evaluate Kant’s claims that ostensibly limit the moral significance of sincerity and communion and reveal that they either fail to capture his own deeper theoretical commitments or obscure key features of relationality in friendship. These interventions and departures serve to capture the potential that systematic attention to reciprocal trust has to generate a contentful conception of moral striving within a broadly Kantian socio-historical vision of morality as a human vocation.

While Kant famously asserts a rigorist prohibition on lying, his views on sincerity and reserve are both much more ambivalent and much more dominated by prudential considerations.26 I will argue that Kant’s defense of reserve undermines the importance of positive commitment to truthfulness, which is clearly served by the cultivation of sincerity.

Recall that Kant uses the term openheartedness to refer to a “candid and sincere disposition.” A lack of openheartedness implies a lack of sincerity, an inability to forgo pretense in social exchanges. There are three basic modes of self-presentation available to us: one may express oneself with candor and sincerity, conceal oneself through some form of deception, or artfully evade self-disclosure and direct declaration of one’s thoughts. Kant puts much stock in the third option, yet he is aware of the fact that declining to share one’s thoughts by maintaining reserve does not always allow one to avoid others’ negative judgments.

Consider that in some contexts silence or a reply of “no comment” on an issue may be readily interpreted as either a refusal to take responsibility for one’s views or as an indication of sympathy with reprehensible views. Kant himself says of the silent man, “if he is asked his opinion of a thing, and says: I have nothing to say, that is as much as if he were to speak against it, for if he thought well of it, he could surely say so. … Silence always gives us away” (C, 27:445). Kant attempts to craft an approach to reserve that is at once sociable, morally permissible and prudent, but it is not clear that he marks out a viable practical strategy.

Because he considers silence to be both unsociable and imprudent, Kant asserts that we must generally ‘speak and pass judgment on

---

26. Kant’s dominant concern with respect to reserve is protection from others’ judgments. In the juridical case, if your trust in others is misplaced your property may be stolen. The parallel in the case of trusting others with your real thoughts is not just loss of social status, for here misplaced trust makes you vulnerable to manipulation for others’ ends and others’ sabotage of your own pursuit of ends. Hence, the problem with Kant’s approach is not that prudential concerns are raised, but rather that they are not balanced with discussion of moral concerns. Judgment is needed regarding when to open oneself up to others; my point is that it is morally important to be able and willing to do so when appropriate and it will never be completely risk free.
would constitute an unstable stance from which to pursue moral self-perfection.

Nevertheless, to some extent, Kant regards the forms of self-misrepresentation addressed by Rousseau as morally neutral reserve, prudential necessities that are not at odds with the cultivation of virtue. He writes,

> In any company we tend to withhold the greater part of our disposition. ... Everyone makes such judgments as are advisable in the circumstances; we are all under constraint, and harbor a mistrust of others, which results in a reserve, whereby we not only cover up our weaknesses, so as not to be ill thought of, but also withhold our opinions. If, however, we can get rid of this constraint, and impart our feelings to the other, then we are fully in communion with him (C, 27:427).

It is desirable to be able to be sincere, but it is not inherently problematic to carefully curate one’s public persona. We might wonder why Kant does not, like Rousseau, see the intentional creation of a gap between reality and appearance as a path to vice. It is plausible to think that there is a meaningful distinction between seeking advantage and avoiding disadvantage. It may seem less dangerous to misrepresent oneself in order to avoid negative judgments, especially unfair negative judgments, than to seek elevated status through self-aggrandizement. Even if this is true, with some qualifications, there are reasons implicit within Kant’s own thought for judging sincerity and the social space for self-disclosure as important means for a sustained commitment to truthfulness and the exercise of self-governance.

**Significantly, in other contexts Kant does frame dissimulation as a serious threat to moral integrity, yet he does not apply this insight within his discussions of friendship. For example, in the Anthropology Kant writes, “It belongs to the basic constitution of the human creature and to the concept of his species to explore the thoughts of others, but to withhold one’s own—a nice quality that does not fail to progress gradually from dissimulation to deception and finally to lying” (A, 7:332).**

27. For a relevant discussion of Kant’s distinction between lies and permissible untruths see Wood (2007), pp. 245–7.
My second argument for the connection between sincerity and a positive commitment to truthfulness centers on the necessity of communion to the expression of moral personhood. As presented in the Doctrine of Virtue, the central role of trust in moral friendship is to fulfill the need for self-disclosure, which is primarily a psychological need, though its fulfillment may serve moral development through increased self-knowledge and improved judgment.\(^3\) Communion with another is intrinsically valuable and meets a deep need, but the relevant question here is whether it is a moral need. Kant never clarifies this point, leaving us to wonder if a candid and sincere disposition is a virtue of friendship that is tangential to other domains of moral striving. In the Vigilantius lecture Kant speaks of the improvement of our judgments as an “inclination converted to a need,” a “pure interest” and “pure goal” that “must lead us to friendship” (\(V, 27:284\)). I suggest that communion may itself count as an inclination converted to a “pure goal,” that is, a moral end. Because mutual sincerity and authentic self-presentation are constitutive of communion, understood as the mutual sharing of thoughts and feelings, sincerity is itself a “pure goal,” one that is by no means peripheral to the pursuit of moral self-perfection.

In his Doctrine of Virtue treatment of lying, Kant asserts that the human being is obligated “to use himself as a natural being in agreement with the declaration (declaratio) of his moral being.” Our capacity to communicate thoughts has a “natural purposiveness” and “inner end,” that we must not subvert (\(MS, 6:429–30\)). A person most fully aligns her use of herself as a natural being with the inner end of communicating thoughts in sincere interpersonal interactions, and this inner end is realized in communion with the trusted friend. Here moral personhood is declared.

---

31. I do not mean to suggest that reciprocal trust is not a moral need; I think it is. However, Kant does not explicitly state this. His presentation of moral friendship in the Doctrine of Virtue seems to treat self-disclosure as a psychological need.

In interpreting the virtue of sincerity in terms of the underlying logic of lying as a vice I do not mean to suggest that insincere self-presentations violate a perfect duty to oneself; rather I intend to draw out the difference between negative and positive expressions of duty. Whereas we express commitment to truthfulness negatively by refraining from telling lies, we express this commitment positively by communicating truthfully. Moral friendship allows friends to be authentic, to not “dissemble” [sich verstellen] in their interpersonal interactions with one another.\(^3\) Of course, some withholding of thoughts and feelings is still necessary, as Kant says friends limit what they share based on the requirements of mutual respect. Yet this level of withholding does not keep us in “prison” because healthy boundaries between people are not oppressive barriers to freedom of expression. Restrictions on self-disclosure motivated by respect as a limiting condition are not deceptive, for here neither anxieties nor ambitions determine choice.\(^3\)

The arguments above both rely on the claim that negative expressions of duty — not violating the ends of virtue — are insufficient for positive commitment to these ends. Intriguingly, moral friendship seems to provide an opening for the positive and direct expression of respect. Although Kant argues that any duty of respect is expressed only negatively and indirectly “through the prohibition of its opposite,” namely disrespect, his description of developed moral friendship represents trusting as an active affirmation of another’s value (\(MS, 6:465\)).\(^3\) Beyond (negative) recognition respect for the person as
such, one friend respects the other (positively) as the individual person she is, in her sincerely communicated and sensitively understood fullness.\textsuperscript{35} The particularity of friendship and its kind of regard can be appreciated by noting that if friendship dissipates recognition respect remains. The respect moral friends have for each other is clearly richer than recognition respect, yet appraisal respect is also not an apt descriptor of the moral bearing of friends.\textsuperscript{36} Appraisal respect is a matter of evaluation, of measuring another in terms of a set standard. Yes, moral friends tend to appraise each other highly, but as I will explain at greater length below in my discussion of being known, relating to another as an object of appraisal is potentially disruptive of communion, whereas positive respect sustains communion.

The relationship between sincerity and positive respect is quite immediate. In moral friendship we respond to the friend herself, as opposed to relating to a self-presentation that is a narrative or appearance produced for others to view. The need to defend ourselves against loss of status through fabricated personas directs us to relate to others from outside of ourselves. But we cannot experience ourselves as valued in the mode of positive respect if we do not make ourselves available for relation with another. One positive expression of duty

love and respect as opposing forces—one drawing people together and the other holding them apart—is relevant here (p. 392). Strictly speaking, this cannot be right, because a maxim of practical love includes respect (p. 396). I agree with Baron that Kant exaggerates the tension between love and respect (p. 406). Lara Denis (2001) argues, along similar lines, that Kant exaggerates the distance needed in friendship.

\textsuperscript{35} Melissa Merritt (2017) argues for a mode of respect along these lines, claiming that “dispositional readiness” to love the other “colours or imbues” the virtuous person’s attitude of respect toward another person “making it what it is as an expression of virtue” (p. 1855). Merritt’s interest in this heightened form of respect lies in its exemplification of attention to the non-interchangeable individuality of the other. While Merritt frames love-imbued-respect as a moral ideal to be pursued in all encounters with others, a similar version of positive respect is actualized by the (less-than-perfectly-virtuous) participants in a moral friendship.

\textsuperscript{36} See Darwall (1977) for further explanation of the distinction between recognition respect and appraisal respect.

I will close my case for the moral significance of sincerity by explaining why sincerity is essential to both negative and positive fulfillments of the duty of self-knowledge. Within the context of moral friendship, the virtue of sincerity relates to self-governance, the “positive command” of virtue, by creating significant barriers to self-deception and deception of others (MS, 6:408). Much of what we learn about ourselves through friendship is not due to the friend enumerating our flaws, but rather a function of our own confessions. Nehamas observes that we speak to friends “often more freely than we speak to ourselves—and reveal aspects of ourselves of which we may be suspicious, unsure or even ignorant.”\textsuperscript{37} Strong bonds of trust draw us out of ourselves, allowing troubling thoughts to come to consciousness and stay in consciousness. Once one has confided in a friend, it is no longer possible to rationalize the uncomfortable issue away or deny its existence. But it is not just that a communicated thought can no longer be ignored or disowned; rather, in sharing our thoughts and feelings we take responsibility for them, making ourselves answerable to both ourselves and another person. The duty of self-knowledge is simultaneously fulfilled in both negative and positive senses, as we avoid self-deception by taking responsibility.

Kant privileges verbal communication and even claims that it is not possible to express one’s feelings without putting them into thoughts (V, 27:677).\textsuperscript{38} But of course we do express feelings, attitudes and even thoughts non-verbally. We communicate meaning without words. One implication of this fact is that we get revealed to others even when we do not wish to be. And our ability to conceal our thoughts, feelings and attitudes decreases when we are known very well by another

\textsuperscript{37} Nehamas (2016), p. 223.

\textsuperscript{38} “For feelings can be disclosed no otherwise, than by the imparting of thoughts; thus we must have an idea of the feeling in advance, and must hence have employed reason, in order to have known it accurately before we share it …” (V, 27:677).
person. In moral friendship a person discloses herself to another both intentionally and unintentionally. Developed friendship makes deception harder, such that if you choose to deceive you need to make a concerted effort, which in turn makes it harder to convince yourself that you are not doing it. In these ways being known itself creates a barrier to self-deception and thus a support to self-governance.

Lastly, the achievement of trust in moral friendship makes it possible to disclose one’s painful experiences of moral failure or inadequacy without fear, creating a pathway to the cultivation of honesty with oneself that potentially goes beyond pre-emptive resistance to the temptation to self-deceive. It becomes possible to learn more, not just to avoid backsliding into self-deception. Freed from the prudential imperative to self-concealment, we have a person with whom to work out more anxious narratives about ourselves and more hopeful narratives about our future development. Knowledge of one’s flaws is necessary to negative moral striving, but often we come to our friends already in possession of that burdensome self-knowledge. In friendship we have a partner in positive moral striving, a person with whom to make use of self-knowledge in pursuit of positive ends.

(2) Being Known
Having observed that being-known alters one’s relation to oneself, I now turn to the broader issue of how communion gives rise to distinctive forms of self-knowledge. In contrast to Kant’s concern with scrutinizing one’s motives and flaws, I emphasize underexplored ways in which being known and knowing oneself in relation to another person may contribute to the project of moral self-cultivation. Here again the goal is to demonstrate what can be gained philosophically by making more of the kind of reciprocal trust definitive of moral friendship than Kant did himself.

The moral value of being known by a friend has traditionally been located in friendship’s contribution to self-knowledge. My sense is that proper appreciation of the moral significance of this aspect of friendship requires a grasp on how self-knowledge gained through relationship with another is different from self-knowledge arrived at through other routes, as well as an explicit explanation of the conditions under which self-knowledge actually facilitates moral improvement. Reflecting on Stanley Cavell’s conception of “passive skepticism,” the worry that one cannot be known, Richard Moran writes,

Taking seriously the question of being known, and the subject’s own ‘say’ over that question, may make the fantasy of ‘peering into the mind of another’ seem not just fantastical or out of reach, but as missing the fact of how the knowledge of another is a matter of relatedness or responsiveness to that person, and not, as it were, simply inspecting them or some other thing “inside” them.39

A person does not observe and inspect her friend and then report her findings. That is not how one typically contributes to a friend’s self-knowledge. A friend communicates with her friend in a way that is responsive to the other’s current self-understanding and to the relationship between the self and other. Unlike pointing out to another person that she has spinach in her teeth, to criticize, question or otherwise respond to a friend’s attitudes, judgments, narratives, actions and plans is to empathically engage with the friend. The friend does not glimpse into the other’s privileged space of awareness, but rather communicates in terms of a shared horizon of meaning. When one is known relationally, it is much more likely that one will experience oneself as known, as understood. Being treated as a something to be inspected and then fixed or repaired damages relationality and thus cuts off communion (even if the inspector is correct in his attribution of flaws). As Kant remarks, “even the fact that his friend observes him and finds fault with him will seem in itself offensive” (MS, 6:470).

Further, communion is not just what gives the other access to the self and sustains the trust needed to take the other’s views of the self seriously. Communion makes possible the transformative experience

of being understood, which is by no means reserved for the morally excellent yet has deep implications for how we conceptualize moral transformation in the context of friendship. Jaxie, the rough and troubled young protagonist of Tim Winton’s novel The Shepard’s Hut captures aspects of the phenomenology of being known with unexpected eloquence in his reflection on Fintan, the hermit-like figure with whom he shelters in the Australian wilderness.

He got me somehow, that’s one thing for certain. Makes sense in a way because we put in bulk time together, those months there was no one else to turn to … there was some hot feeling he give off once you knew him. It was like you were standing too close to the stove or coming over with a fever. … It’s a dangerous feeling getting noticed, being wanted. Getting seen deep and proper …. It’s like being took over. And your whole skin hurts like you suddenly grew two sizes in a minute.40

There is a heat in being known that matters just as much, if not more, than the light that might be shed on the self by a perceptive observer. For people more familiar with intimacy than Jaxie, being known will likely be experienced more often as a pleasant warmth than an uncomfortable flame, and a friend’s empathy will likely feel less like an invasion. Still, Jaxie’s sense of expanding beyond his boundaries is germane, as it indicates that in communion we are in fact different than we are in isolation.

To take communion seriously is to move away from a model of the individual striving for moral improvement in private, according to which the friend issues criticisms with which one skulks away to contemplate.41 Shifting focus to a communal model of moral striving improves Kant’s account because it explains how friends can counsel each other without trust-compromising threats to mutual respect. We might say that Kant has half the picture, the part of the picture that poses the biggest threat to appraisal respect, namely, the direct attribution of a flaw. A moral friendship can handle this kind of exchange because it is not the dominant mode in which friends know each other. Relational knowing through which friends collaboratively problem solve, contextualize each other’s flaws, help each other consider options for thought and action, and clarify moral commitments is the background upon which direct attribution of flaws can be productive. My sense is that this background is more central to striving for moral perfection than the kind of self-knowledge usually placed in the foreground of philosophical discussion of friendship and moral improvement.

When my friend points out my flaws, the target of criticism may well be my treatment of or mode of relating to the friend. (People who are friends with their spouse know this well.) Kant’s idealization of friendship keeps this out of view, as he associates conflict with inferior forms of friendship, remarking that “the rabble fight and make up” (MS, 6:471). Yet trust is what allows friends to navigate conflicts. We need a model for resolving conflicts and threats to mutual respect, and moral friendship is best suited to provide it, as here trust and respect for principle is strongest. In sharing their real thoughts and feelings, including moral assessments and moral confusions, friends refine their interpretations of principle and explore the implications of these interpretations. Friendship challenges the agent more concretely when struggles to put principle into practice or to enact self-governance cause tension within the relationship. In relationships of well-developed reciprocal trust, we are able to work through problematic relational dynamics, misunderstandings and disagreements. If what we say is ill received, we have a way to fix it — through more communication. Our commitment to the end of moral self-perfection is also a commitment to the friend. We work to develop shared moral meanings where they do not...

40. My discussion here might be read as a compliment to Rudolf Makkreel’s (2014) analysis of the ways in which the Kantian agent might compensate for the limits of introspective self-observation. In a similar vein, Laura Papish (2018) writes, “Self-knowledge is not in its essence a kind of internal archaeological dig but an ongoing exercise in self-interpretation” (p. 172).
already exist and we strive to correct failures of self-governance for the sake of the relationship.

Considerable moral growth can be accomplished through resolving miscommunications, conflicts, painful offenses, etc. It is not just that “uncultivated people” need quarrels in order to savor the sweetness of being united in reconciliation” (MS, 6:471). Conflicts arise between people who are closely bound, especially between people who live together, that would never arise between co-workers, or even friends located at a distance from each other. Is it better to live without these conflicts? Those who arrange to live alone and not spend too much concentrated time with others may need to make fewer compromises and can minimize being revealed in action. However, there is value in being known. There is intrinsic value in being understood, in experiencing communion of thought and feeling, but being known also creates opportunities for the development of more effective self-governance. Taken together, the many ways in which (1) sincerity and (2) being known contribute to (3) self-governance establish the importance of moral friendship, and the virtues that sustain it, to the fulfillment of the duty of moral self-perfection.

42. Emily McRae (2017) directly addresses the challenge posed by such mutual exposure of minor personal faults and failings. She writes, “No matter how lovable our beloved, the intimacy of love ensures that we will eventually witness some of her ugliness. And, of course, no matter how committed we may be to self-improvement, it is nearly impossible for our loved ones to escape having to deal with our own ugliness” (p. 337).

43. Though I focus here on being known, we might think about the inevitable other side of living in relationship, which is knowing the other and responding to their mistakes. A line from D. J. Waldie’s Holy Land (1996) captures this idea with uncanny poignance: “The greatest loss in living deliberately alone is not having anyone to forgive” (p. 96).

44. As Eric E. Wilson (2008) has explained, the distinction between heteronomy and autonomy (between what is external and what is internal to the self) is normative, not metaphysical (p. 372). There is no need to worry that my emphasis on communion gives the friend too much of a foothold in the other’s mental space.

Conclusion

I have argued that the overlap between the duty of friendship and the duty of moral self-perfection extends beyond the role of friendship in self-knowledge and refinement of judgment. My attention to the nature of trust, as well as to Kant’s socio-moral psychology, reveals a more fundamental relationship between reciprocal trust and moral self-cultivation than has previously been recognized. I have looked outside of Kant’s texts in order to develop an account of moral friendship that remains true to his central theoretical commitments. The idea that moral perfection cannot be effectively pursued by a single, isolated individual is Kant’s own clearly stated claim. If the pursuit of moral self-perfection is not to be a Sisyphean task, this or some other story about reciprocal trust must support a Kantian account of moral striving.

I hope to have also shown that we need not be so partisan in our approaches to the philosophy of friendship that we can take interest only in the Elizabeth and Darcys of the world or only in the Thelma and Louises. There are elements of moral friendships in non-moral friendships and there are dimensions of striving that traverse the boundaries of different forms of relationship. The point of a conception of moral friendship is to give us a model for thinking about how commitment to moral principle might first enable and later be transformed by trusting engagement with another person. It also helps explain why friendship is the greatest of human goods, in those cases when this is clearly so.

45. I am grateful to the audience at the 2018 Central APA NAKS Biennial, and especially to my commenter at the latter venue, Franey Russel, for feedback on early versions of this project. Conversations with Drew Leder, Emily McRae and Ellen Graf informed my thinking about being known and my selection of examples. I thank the following members of the DC-Baltimore Kant Workshop for helpful feedback at later stages of the project: Alexander Englert, Tim Jankowiak, Michael Nance, Laura Papish and Krista Thomason. Finally, I thank two referees for this journal for their very useful comments.
References

Works by Immanuel Kant:

Note: English translations of passages from the Metaphysics of Morals (1785) are from Immanuel Kant: Practical Philosophy, Mary J. Gregor trans. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press, 1996. Because all of the translations reference the volume and page numbers of the Akademie edition of Kant’s collected works, I include only the Akademie edition page numbers in order to simplify the citations. I use the following abbreviations of Kant’s titles: CPR = Critique of Pure Reason; MS = Metaphysics of Morals; R = Religion within the Boundaries of Mere Reason; V = Vigilantius (in Lectures on Ethics); C = Collins (in Lectures on Ethics); A = Anthropology from a Pragmatic Point of View.


Other Sources:


