REGULAR PAPER

Virtue Ethics Must be Self-Effacing to be Normatively Significant

Scott Woodcock[1] (ORCID)

For an ethical theory to provide meaningful normative guidance, it ought to equip agents with more than just abstract rules and principles; it ought to also supply agents with useful practical advice about how to live their lives in a complex world of values and obligations. Appreciation of this role for ethical theory is one of the driving concerns of those who favour contemporary formulations of virtue ethics. Supporters of virtue ethics regularly emphasize that rival ethical theories, like consequentialism and rule-based deontology, fail to adequately account for the *lived experience* of agents seeking to incorporate normative prescriptions into the nuances of their daily lives. In particular, many of those who support virtue ethics assert that their view is immune to the "moral schizophrenia" that Michael Stocker diagnoses for ethical theories that require self-effacement in the moral psychology of agents to capture goods like friendship, i.e. goods that cannot be pursued for the sake of abstract justifications without losing the special value they contribute to our lives.[1] Stocker's thesis has been influential, and the presumption that virtue ethics is not vulnerable to self-effacement has led many to conclude that virtue ethics maintains at least one advantage over its contemporary rivals.

In this paper, I argue that this presumption is unjustified and that virtue ethics is just as vulnerable to self-effacement as rival ethical theories. Thus, my argument is that Simon Keller is correct in his assessment that the options available for virtue ethics to avoid self-effacement, or at least diffuse its problematic features, are options that are equally available to rival ethical theories.[2] I begin with a summary of the literature preceding Keller's thesis that virtue ethics is self-effacing. I then

---

[1] Michael Stocker, "The Schizophrenia of Modern Ethical Theories," *The Journal of Philosophy*, vol. 73, no. 14 (1976): 453-466.

[2] Simon Keller, "Virtue Ethics is Self-Effacing," *Australasian Journal of Philosophy* vol. 85, no. 2 (2007): 221 - 231.

✉ Scott Woodcock
   woodcock@uvic.ca

[1] Department of Philosophy, University of Victoria, STN CSC, PO Box 3045, Victoria, BC V8W 3P4, Canada

discuss three noteworthy attempts to undermine Keller's arguments, and I conclude that each of them fails in a way that reveals something interesting about the features of ethical theories that trigger a need for self-effacement. Most notably, I conclude that virtue ethics cannot escape self-effacement if it retains the ability to provide agents with substantive normative guidance. After presenting this argument, I briefly discuss the implications for virtue ethics if the argument is sound, and I suggest that advocates of virtue ethics remain open to the possibility that self-effacement is a tolerable psychological challenge.

1 A Brief History of Self-Effacement

Stocker's moral schizophrenia diagnosis derives much of its appeal from his example of a friend named Smith who comes to visit you in the hospital.[3] In the example you have been recovering from a long illness when Smith arrives to boost your spirits. You are initially pleased that he has taken the time to cheer you up; however, you discover that he was motivated to visit you from a sense of moral obligation. Whether from devotion to communism, the math of utilitarianism or a strategic calculation of his own long-term self-interest, Smith has determined that he has a duty to visit you and this is what motivated his visit. For most readers, the example produces a sense of alienation and the judgment that his visit fails as a genuine act of friendship. It reveals that some goods cannot be directly pursued via motives that explicitly refer to the *justifications* for one's actions. To retain their authenticity, goods like friendship must be realized by agents who are motivated by the immediate details of their situations, e.g. the mere fact that a friend is ill and in need of companionship.[4]

The initial problem, then, is that ethical theories grounded in general, impersonal duties seem unable to capture an important class of ethical goods, yet Stocker quickly acknowledges that it is possible for these theories to capture goods like friendship if they separate the *motives* that drive agents to act justifiably from the essential *reasons* that justify their actions. The key to his argument is that this type of separation of reasons from motives in the minds of human agents precipitates a "malady of the spirit" that he refers to as moral schizophrenia.[5] Even if it is *possible* for so-called modern ethical theories to capture goods like friendship, to do so these theories must employ a self-effacing strategy that divides agents' justifications for their actions from the immediate motivations that drive them to perform the actions. According to Stocker, this separation of our reasons from our motives is so disruptive that it constitutes a *reductio* for any ethical theory that requires such self-effacing measures as part of its moral psychology.

---

[3] Stocker (1976), p. 462.

[4] Friendship is therefore a good that is "calculatively elusive" in that it "it cannot be attained under the calculative choice of action." See: Philip Pettit and Geoffrey Brennan, "Restrictive Consequentialism," *Australasian Journal of Philosophy* vol. 64, no. 4 (1986): 438 – 455, p. 442.

[5] Stocker (1976), p. 454.

As I have argued elsewhere, Stocker's moral schizophrenia is much more complex than it initially appears.[6] Nevertheless, many share Stocker's conviction that self-effacement is not compatible with a healthy human psychology, and many are similarly persuaded that the need for self-effacement is unique to modern, duty-based ethical theories rather than revitalizations of ancient virtue ethics. The most prominent exception here is Thomas Hurka, who argues that virtue ethics is similarly vulnerable to the alienation that creates a need for self-effacement.[7] In fact, Hurka claims that virtue ethics may even be *more* susceptible to self-effacement compared to other ethical theories, like consequentialism, since these theories are only contingently self-defeating when it comes to situations that require agents to implement their ethical theory via indirect methods. By contrast, according to Hurka, virtue ethics *necessarily* leads to moral self-indulgence if its agents are motivated to promote their own flourishing or the virtuous status of their actions. In reply, Julia Annas claims that virtue ethics is only vulnerable to Hurka's claim if it is uncharitably interpreted as a form of flourishing *egoism*.[8] She presents an interpretation of virtue ethics according to which a virtuous agent is not moved by self-indulgent thoughts about promoting her own flourishing because flourishing is not viewed by the agent as a result that is instrumentally achieved and thus specified independently of the virtues. Instead, Annas views flourishing as a formally defined concept *constituted* by virtuous activity and directed toward the ends of others as much as one's own.[9] This avoids the charge that virtue ethics necessarily involves egoism or self-indulgent motives, and Annas therefore considers the view to be safely immunized from the self-effacement other theories require to motivate their agents without justifications that are toxic to the pursuit of goods like friendship.

It is at this point that Keller presents a new argument for the position that virtue ethics is self-effacing. Without relying on the claim that virtue ethics is foundationally egoistic, Keller offers a simple example to illustrate that virtue ethics, like other theories, ought to sometimes recommend that agents avoid being directly motivated by the reasons that explain why actions are ethically justifiable. He imagines three agents who each perform the same virtuous act of helping a family of hikers trying to set up a campsite in stormy conditions.[10] Arthur is motivated by the immediate

---

[6] Scott Woodcock, "Moral Schizophrenia and the Paradox of Friendship," *Utilitas* vol. 22, no. 1 (2010): 1 – 25.

[7] Thomas Hurka, *Virtue, Vice, and Value* (New York: Oxford University Press, 2001), pp. 246-49.

[8] Julia Annas, "Virtue Ethics and the Charge of Egoism," in *Morality and Self-Interest*, Paul Bloomfield ed. (Oxford: Oxford University Press, 2007), pp. 205-221. Other attempts to circumvent egoistic interpretations of virtue ethics include: Christine Swanton, "Virtue Ethics and the Problem of Indirection: A Pluralistic Value-Centred Approach." *Utilitas* vol. 9, no .2 (1997): 167-181; Dennis McKerlie, "Aristotle and Egoism," *The Southern Journal of Philosophy* vol. 36, no. 4 (1998): 531-555; and Christopher Toner, "Virtue Ethics and the Nature and Forms of Egoism," *Journal of Philosophical Research* vol. 35 (2010): 275-303. For a reply, see: Tom Peter Stephen Angier, "Aristotle and the Charge of Egoism," *The Journal of Value Inquiry* vol. 52, no. 4 (2018): 457-475.

[9] For a similar view of virtuous actions, see: Jennifer Whiting, "Eudaimonia, External Results, and Choosing Virtuous Actions for Themselves," *Philosophy and Phenomenological Research* vol. 65, no. 2 (2002): 270-290.

[10] Keller (2007), pp. 225-26.

details of the situation, e.g. the hikers are cold, wet, and tired. Benjamin is motivated by his desire to act generously and an assessment that helping the hikers constitutes a generous act. Finally, Christine is motivated by a desire to do what the fully virtuous person would do in this situation. The example illustrates that even if we work with a circumspect view of the reasons that virtue ethics puts forward to justify the actions it endorses, we still find that the most praiseworthy agent seems to be someone like Arthur – someone not motivated by the reasons why his acts are virtuous. Keller anticipates that advocates of virtue ethics will reply by claiming that his example rests on a misunderstanding of *de re* versus *de dicto* interpretations of the virtue-based recommendation to act "as the virtuous person would act".[11] However, he points out that this strategy is not uniquely associated with virtue ethics – it is instead one that can be employed by rival theories to similarly avoid the charge of self-effacement. Thus, Keller concludes that virtue ethics faces a dilemma: either it is vulnerable to the same alienation that leads to self-effacing measures, or there are options for it to avoid self-effacement that are not unique to virtue ethics and therefore give it no advantage over other theories.

2 Clark's Developmental Proposal

The most recent attempt to undermine Keller's conclusion comes from Justin C. Clark.[12] Clark expands on some important features of Annas' eudaimonistic virtue ethics in an effort to claim that its agents are not burdened by a problematic requirement to hide their justifications for virtuous actions from the motives that lead them to undertake these actions. He begins with some important distinctions. First, he describes two different levels of deliberation for agents considering how to live their lives: a *global* level of deliberation, at which agents decide how to shape their overall lives, and a *local* level at which agents determine how to act in particular circumstances.[13] One's own flourishing may enter into deliberations when considering the values that guide the direction of one's life, but eudaimonistic virtue ethics need not require that agents think about flourishing when they are deliberating about particular actions in their lives. Next, Clark distinguishes between substantive and explanatory accounts of right action: a *substantive* account identifies common features of right actions and allows agents to identify these actions, whereas an *explanatory* account offers insight into the reasons why actions are right and in what rightness itself consists.[14] It is important to note, for Clark, that the criterion of right action stating that an act is right iff it is what a virtuous person would characteristically do in the relevant circumstances is a substantive account of right action – it is not an

---

[11] The inspiration for this reply is Bernard Williams, "Acting as the Virtuous Person Acts," in *Aristotle and Moral Realism*, R. A. Heinaman ed. (New York: Routledge, 1995), pp. 13-33.

[12] "Eudaimonistic Virtue Ethics and Self-Effacement," *The Journal of Value Inquiry* vol. 50, no. 3 (2016): 507-524.

[13] Ibid. pp. 511-512.

[14] Ibid. p. 513.

account that explains why this action ought to be performed. Thus, it is consistent with a substantive criterion of right action for the motivation of actions to be indefinite and for agents to perform actions *consistent* with virtue from motives other than those specified by the criterion.

With these distinctions in place to help inform his argument, Clark then emphasizes the development of practical wisdom in virtuous agents. He emphasizes Aristotle's key insight that human agents acquire the virtues through a slow process of emulating others and repeatedly engaging in virtuous activity.[15] This process continues until an agent's character is shaped so that she perceives the nuances of particular circumstances and recognizes how to balance the intersecting recommendations of multiple virtues that may apply. Practical wisdom is in this respect a special, comprehensive virtue that enables agents to absorb the fine details of their immediate circumstances and to respond appropriately to such circumstances without having to engage in explicit calculations regarding their options. Agents exhibiting practical wisdom understand the significance of their choices – they are not operating as mere automata – and justifications for their actions can be recovered if they are required to provide reasons for their choices.[16] Yet Clarke points out that the developmental formation of practical wisdom allows an explanatory account of right action for virtue ethics to be *transparent* in the sense that agents can become capable of seeing through the account to the immediate details of the situations they encounter. Consequently, well-developed agents can be motivated to act on reasons that direct them to perform actions for their own sake. This allows criteria of right action that would otherwise cause alienation, e.g. an action is right iff it promotes the agent's flourishing, to dissolve into criteria of right action that can serve as non-alienating motives in particular circumstances, e.g. an action is right iff it is performed for reasons appropriately perceived as applicable to a given situation (and unsurpassed by any other morally relevant considerations).[17]

Clarke is certainly correct to focus on *phronesis* as part of the solution to Stocker's moral schizophrenia challenge. Practical wisdom has an essential role in the structure of virtue ethics, and the development of this virtue is especially helpful for explaining how agents can apply the recommendations of ethical theory to the details of their lived experience. However, as a reply to Keller and the risk of self-effacement, Clarke's proposal is unsuccessful for two noteworthy reasons. First, as Keller foresees, there is nothing prohibiting ethical theories other than virtue ethics from similarly relying on the development of practical wisdom to guide their agents. The substance of what each theory recommends will differ, but nothing prohibits

---

[15] *Nicomachean Ethics* (Oxford: Oxford University Press, 2002). See also: Richard Kraut, "Aristotle on Becoming Good: Habituation, Reflection, and Perception," in *The Oxford Handbook of Aristotle*, Christopher Shields ed. (Oxford: Oxford University Press, 2012), pp. 529-557.

[16] Clark op. cit., p. 514. Note, however, that agents with nuanced capacities for practical wisdom may not be able to give justifications for their choices in codifiable terms that inexperienced agents can understand. See Rosalind Hursthouse, "What Does the Aristotelian Phronimos Know?" in *Perfecting Virtue: New Essays on Kantian Ethics and Virtue Ethics*, Lawrence Jost ed. (Cambridge, Cambridge University Press: 2011), pp. 38-57.

[17] Clarke op. cit., p. 516.

so-called modern ethical theories from emphasizing how important it is for agents to undergo a careful process of moral education that enables them to recognize the details of their immediate situations and to respond to these details without having to explicitly consult the foundational values that inform their explanatory criteria of right action. In fact, this is precisely what contemporary versions of consequentialism propose to address the problem of alienation from goods like friendship: they emphasize the development of robust decision-making capacities in agents equipped with long-term dispositions to recognize, and be motivated by, the morally salient details in their specific circumstances.[18] Kantian ethics, as well, is not restricted to proposing agents who develop nothing more than a general disposition to apply the categorical imperative at every available opportunity.[19]

Thus, Clarke's proposal to connect the development of practical wisdom to the debate over self-effacement is both (a) helpful insofar as it identifies something interesting about what is required for agents to be governed by normative principles without losing their capacity to be motivated by the relevant features of their specific circumstances, and (b) inadequate as a reply to Keller because the development of practical wisdom is not restricted to agents who endorse virtue ethics rather than other ethical theories. In effect, Clarke emphasizes the developmental details that allow agents to flourish in a *de re* sense of acting as a virtuous person would act – a proposal that avoids the alienation that arises if agents are motivated by the same standard if it is interpreted as a *de dicto* criterion of right action. This is a sensible strategy, but it is explicitly discussed by Keller who notes that it has no essential connection to virtue ethics and therefore moves the debate over self-effacement no further forward.[20]

---

[18] For example, the rule-consequentialism defended by Brad Hooker takes into account agents' dispositions in its calculations of an optimific set of rules to be internalized by a majority of agents over successive generations, and the dispositions recommended by this optimific set of rules presumably include those that allow agents to capture special goods like friendship that must be pursued from authentic personal motives. See: Brad Hooker, *Ideal Code, Real World: a Rule-Consequentialist Theory of Morality* (Oxford: Clarendon Press, 2000). Other types of indirect consequentialism and sophisticated versions of act-consequentialism are also structured so agents develop finely-tuned dispositions to act directly for the sake of calculatively elusive goods without constantly reverting to direct maximization. See: Peter Railton, "Alienation, Consequentialism, and the Demands of Morality," *Philosophy & Public Affairs* vol. 13, no. 2 (1984): 134-171; Frank Jackson, "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection," *Ethics* vol. 101, no. 3 (1991): 461-482; Elizabeth Ashford, "Utilitarianism, Integrity, and Partiality," *The Journal of Philosophy* vol. 97, no. 8 (2000): 421-439; and Julia Driver, *Consequentialism* (New York: Routledge, 2012), pp. 102-13.

[19] See, for example: Nancy Sherman, *Making a Necessity of Virtue* (Cambridge: Cambridge University Press, 1997), ch. 7; Sergio Tenenbaum, "Friendship and the Law of Reason: Baier and Kant on Love and Principles," in *Persons and Passions: Essays in Honor Of Annette Baier*, Joyce Jenkins, Jennifer Whiting, and Christopher Williams, eds. (Notre Dame, Ind: University of Notre Dame Press, 2005), pp. 250-280; Barbara Herman, *Moral Literacy* (New York: Harvard University Press, 2007); Marcia Baron, "Virtue Ethics, Kantian Ethics, and the 'One Thought Too Many' Objection," in *Kant's Ethics of Virtue*, Monika Betzler ed. (New York: Walter de Gruyter, 2008), pp. 245-278; and. "Kantian Moral Maturity and the Cultivation of Character," in *Kant on Emotion and Value*, Alix Cohen ed. (London: Palgrave Macmillan, 2014), pp. 69-87.

[20] Op. cit., p. 230.

The second reason Clarke's proposal fails as a reply to Keller is that it only allows virtue ethics to avoid self-effacement if we presume as our representative agent someone who is the *perfect* instantiation of the *phronomoi*, i.e. someone who has developed a perfect capacity for practical wisdom and who will therefore perceive every morally salient detail in every possible situation and know how to balance any competing moral considerations in just the right way so that she never needs to step back and reflect on the foundational values that justify her ethical obligations. Consider, for example, Annas' description of the process of virtue acquisition:

> A beginner in virtue will have to try explicitly to become a virtuous person, and to do so by doing virtuous actions; his deliberations will include such thoughts as that so and so is what a virtuous person would do, or what virtue requires. This is, indeed, how he guides his own deliberations. The truly virtuous person, however, will not explicitly think about, for example, being brave or performing a brave action. Rather, he will, as a result of experience, reflection, and habituation, simply respond to the situation, thinking that these people in danger need help, without explicit thoughts of bravery entering his deliberations. Thoughts about bravery, or the virtuous person, are no longer needed.[21]

The ideal agent proposed by Annas and Clarke avoids a continuing need for psychological self-effacement because the threat of alienation from one's immediate circumstances never arises when this agent is already perfectly disposed to act for the sake of her friends in *just the right way* as each situation requires. She will visit you in the hospital simply because she cares (just the right amount) for you for your own sake, so Stocker's initial alienation problem never gains purchase. Yet it is worth noting how many important moral considerations must be balanced under the hood, so to speak, by this ideal agent. She will recognize without reflection that she should not ignore cries for help from a nearby assault victim, for example, just because she is in a hurry to see you before visiting hours are over. She will simply know not to visit for too long in a way that could be ingratiating. She will know – again, without any explicit reflection – just how frequently she ought to visit without undermining her professional obligation to finish a research proposal she has promised to complete. The list could go on and on. Essentially, she will recognize at each and every moment that her time could not possibly be spent in any way that better instantiates the full spectrum of virtues compared to what her developed capacity for practical reason has already motivated her to carry out.

This notion of a perfectly virtuous agent is certainly inspiring, and it is perhaps a useful ideal for which agents ought to strive to achieve. Nevertheless, it cannot serve as a plausible representation of how virtue ethics avoids self-effacement for human agents acting in realistic circumstances. It would display an ironic vice of hubris to think that even the wisest and most experienced of us should not sometimes reflect

---

[21] Annas (2007), p. 212.

on the overall balance of our various normative commitments.[22] As human beings we are agents with limited epistemic abilities, and we must therefore repeatedly reflect on whether we are exhibiting the most admirable sum of virtuous dispositions over the course of our lives. Feminist revisions to the presumed substance of the virtues and results from cognitive science exposing our epistemic limitations help to make this point particularly vivid.[23] But the mere fact that we are imperfect beings, rather than gods or archangels, is sufficient to establish that an ideal embodiment of practical wisdom is not a fair test case for determining how virtue ethics provides normative guidance to ordinary human agents. Thus, the debate over self-effacement cannot be settled by invoking agents with such flawless capacities for practical wisdom that they need never reflect and deliberate about how to best implement the virtues in their unique circumstances.

3 Martinez's Appeal to Indirection

So far we have seen that practical wisdom is an indispensable feature of virtue acquisition, yet it is a feature that fails to give virtue ethics any special advantage in the debate regarding self-effacement. If Stocker's hospital case remains the flagship illustration of how ethical theories require self-effacement to avoid alienation, proposing a Smith so practically wise that he need never reflect on whether his visit is justified (and potentially be moved by his reflections) is no fair method of separating one ethical theory from another. Any reasonable ethical theory can propose a super-Smith who exhibits the core values of the theory without risking the alienation that arises if agents reflect on the justifications for their decisions. The trick is to find a way for one's theory to allow agents to periodically engage in this type of reflection without tainting the authenticity of agents' special commitments in their lived experiences.[24]

---

[22] For a strong defense of this point, see: Ron Aboodi, "One Thought Too Few: Where De Dicto Moral Motivation is Necessary," *Ethical Theory and Moral Practice* vol. 20, no. 2 (2017): 223-237.

[23] A small sample of feminist perspectives on virtue ethics: Susan Moller Okin, "Feminism, Moral Development, and the Virtues," in *How Should One Live?* Roger Crisp, ed. (Oxford: Clarendon Press, 1996), pp. 211-229; Lisa Tessman, *Burdened Virtues: Virtue Ethics for Liberatory Struggles* (Oxford: Oxford University Press, 2005); Sandrine Berges, *A Feminist Perspective on Virtue Ethics* (New York: Palgrave Macmillan, 2015). As for results from cognitive science, I am not here referring to literature on the situationist critique of virtue ethics. My point is only that if one is aware of the many ways in which humans are susceptible to various types of cognitive and implicit bias, then they ought to reflect on their implementations of the virtues in particular contexts. To presume that one has already perfectly internalized the virtues so reflection is no longer required would be to ignore literature like: Daniel Kahneman, *Thinking, Fast and Slow* (New York, Farrar, Straus and Giroux, 2011); Alex Madva, "Virtue, Social Knowledge, and Implicit Bias," in *Implicit Bias and Philosophy, Vol. 1*, Jennifer Saul & Michael Brownstein, eds. (Oxford: Oxford University Press, 2016), pp. 191-215; and Dennis Whitcomb et al., "Intellectual Humility: Owning Our Limitations," *Philosophy and Phenomenological Research* vol. 94, no. 3 (2017): 509-539.

[24] The tension here is nicely articulated by Annas in her discussion of the similarity between virtue acquisition and the acquisition of practical skills like those of dancers, carpenters and musicians: "Neither is merely mindless habit; neither is a constantly busy conscious presence in our activities." ("Applying Virtue to Ethics," *Journal of Applied Philosophy* vol. 32, no. 1 (2015): 1-14, p. 4.) A deli-

This leads to the next noteworthy reply to Keller. Joel A. Martinez argues that virtue ethics employs a *harmless* type of indirection in the way reflection is incorporated into action-guiding recommendations for moral agents.[25] By way of contrast, Martinez claims that modern ethical theories employ a different type of indirection that is *pathologically* self-effacing; hence, virtue ethics retains an advantage over its rivals when it comes to moral motivation. The key to his argument is this contrast between mere indirection and self-effacement. Mere indirection, according to Martinez, is exhibited when agents guided by an ethical theory ought to avoid focussing on the background justifications for their actions in order to successfully achieve their aims. For example, Sophia will more successfully act on the virtue of beneficence if she attends to the particular details of others in need of aid than if she is explicitly motivated by her ethical theory's criterion of right action. This much is consistent with Keller's example of helping hikers set up camp, but what is crucial, according to Martinez, is that indirection occurs when *no conflict* exists between agents' criterion of right action and the motives recommended for them to successfully fulfil this criterion. A fully expressed version of virtue ethics, Martinez argues, offers a detailed account of the right-making features of actions so that the particular reasons motivating agents are *consistent with* more schematic descriptions of these reasons, e.g. acting in a way that exhibits beneficence, or acting as a virtuous person would act. In this respect, a fully expressed account of virtue ethics can "make perspicuous the possibility that motives and reasons may be distinct without thereby being in conflict."[26]

The description of indirection in virtue ethics that Martinez provides is compelling, and it captures an important feature of applying criteria of right action to our lived experience: we don't want agents to reflect and be directly motivated by their abstract criteria of right action, yet it is normally unproblematic to have reasons derived from these criteria available if agents are called to reflect on their decisions.[27] Problems arise, however, when Martinez asserts that modern ethical theories require self-effacement rather than mere indirection. These theories are said to require a uniquely harmful type of indirection because the reasons they provide to

agents via criteria of right action are *in conflict with* the motives the theories recommend for avoiding alienation in specific circumstances. Though it may be possible for agents to operate with this self-effacement of fundamental principles, Martinez (following Stocker) argues that tension between motives and reasons is deeply objectionable for what might otherwise be an acceptable type of indirection in an agent's moral psychology.[28]

There are two ways in which this argument from Martinez is unsuccessful. First, though it may seem plausible for philosophers to diagnose tension between levels in an agent's moral psychology as harmful, it is conspicuous that no empirical evidence is ever provided to support this diagnosis. Proponents of virtue ethics take it to be self-evident that it is injurious to agents if they are sometimes required to compartmentalize certain aspects of their thinking from their immediate motives. Ultimately, however, this is an empirical claim about human psychology, and one in need of support for us to conclude that the tension in question is worse than other kinds of curious mental strategies, e.g. setting aside self-preservation for the sake of courage, or trying not to focus on transgressions for the sake of forgiveness. Without this support, the expectation that any kind of tension in human moral psychology is harmful looks more like an aesthetic preference than an argument that avoids begging important questions about theory implementation. Moreover, the burden of proof here falls squarely on Martinez to make the case that self-effacement is specially detrimental compared to mere indirection.[29]

Nevertheless, even if one agrees with Martinez that the harmfulness of self-effacement is self-evident, there is another reason why his argument against modern ethical theories does not succeed. This second reason is that he provides insufficient support for his assertion that modern ethical theories provide agents with justifying reasons that conflict with the motives they recommend for agents to successfully achieve their aims. Martinez claims that modern theories require self-effacement whereas virtue ethics merely requires indirection, but this is not obvious. If one subscribes to a Kantian variant of deontology, for example, why think that conflict exists between the motives that lead you to visit a friend in the hospital (because you care about them for their own sake) and the justificatory reason that this is a case of treating others as ends in themselves?[30] It would be alienating to visit a friend with the explicit motive that you are aiming to fulfill the categorical imperative, but this is equivalent to what we have already noted about acting to help others with an

---

[28] In Pettit and Brennan's terminology, the argument is that modern ethical theories fail because certain goods are more than just calculatively elusive – they are also calculatively *vulnerable* for these theories because the value of the goods evaporates if they are monitored by justificatory reasons that conflict with agents' authentic motives.

[29] Ben Eggleston makes a related point when he argues that if the publicity condition, which rules out self-effacing theories, is to have any useful role to play in the evaluation of moral theories then it must work with facts about the actual world rather than be interpreted as a point of principle. "Rejecting The Publicity Condition: The Inevitability of Esoteric Morality," *The Philosophical Quarterly* vol. 63, no. 250 (2013): 29-57, pp. 54-55.

[30] On this point, see Cynthia Stark, "Decision Procedures, Standards of Rightness and Impartiality," *Noûs*, vol. 31, no. 4 (1997): 478-95, p. 484.

explicit motive to act as the virtuous person would act. In both cases we have justifying reasons that should not serve as direct motives, yet these reasons are available if agents need to reflect on the status of the more specific motives that normally guide their decisions. It is not clear why the deontological case must be viewed as asymmetrical to Martinez's own virtue ethics examples, since viewing one's friends as ends in themselves is hardly *in conflict* with viewing them as unique individuals. Similarly, it is not clear why acting from motives like immediate compassion for others is necessarily in conflict with consequentialist reasons to promote the good.[31] Martinez takes it as given that modern ethical theories operate on tainted criteria of right action – criteria that cannot be reconciled with authentic motives for acting in particular circumstances – but this presumption requires support to play such an important role in his argument. Otherwise, Martinez is *starting out* with an unfavourable view of modern theories before issues of moral psychology even come into play.

The underlying basis for this premise, I suspect, is a claim that modern ethical theories supply agents with reasons that conflict with their immediate motives (to, say, act directly for the sake of their friends) because the reasons reflect abstract, impartial considerations that a modern ethical theory will generate from some comprehensive perspective. In other words, I think philosophers like Martinez presume that modern ethical theories necessarily generate reasons that are toxic to agents' particular commitments because the reasons derive from an overall, often universal, understanding of ethical obligation – an understanding that inevitably leads to alienation because it articulates an impersonal, *sum total* of agents' obligations that cannot capture their particular commitments. If this is correct, however, then Martinez runs into a new problem, because it is not clear that virtue ethics can avoid the same potential for alienation created by the contrast between overall and particular perspectives. Although it is not commonly emphasized, virtue ethics also provides agents with a comprehensive view of how all the normative recommendations of its manifest variety of virtues fit together. If this is acknowledged, it becomes evident that virtue ethics faces the same challenge of alienation for agents who reflect on the *sum total* of their obligations (e.g. those grounded in justice, honour, prudence, etc.)

---

[31] This is the point of Railton's memorable case of Juan, who acts from direct concern for his wife Linda though he can, if prompted to reflect, reconcile his concern with consequentialist justifications. (Op. cit., p. 150) The example is, of course, not decisive. Some dispute the possibility that relationships can be authentic for consequentialists, e.g. William H. Wilcox, "Egoists, Consequentialists, and Their Friends," *Philosophy and Public Affairs* vol. 16, no. 1 (1987): 73-84; Neera Kapur Badhwar, "Why It Is Wrong to be Always Guided by the Best: Consequentialism and Friendship," *Ethics* vol. 101, no. 3 (1991): 483-504; and Dean Cocking & Justin Oakley, "Indirect Consequentialism, Friendship, and the Problem of Alienation," *Ethics* vol. 106, no. 1 (1995): 86-111. Consequentialists, for their part, have responded to this criticism, e.g. Alastair Norcross, "Consequentialism and Commitment," *Pacific Philosophical Quarterly* vol. 78, no. 4 (1997): 380-403; Elinor Mason, "Can an Indirect Consequentialist be a Real Friend?" *Ethics* vol. 108, no. 2 (1998): 386-393; and Scott Woodcock, ""When Will your Consequentialist Friend Abandon You for the Greater Good?" *Journal of Ethics & Social Philosophy* vol. 4, no. 2 (2010): 1-23. Given this outstanding debate, the burden of proof remains on Martinez to explain why consequentialist justificatory reasons inevitably conflict with authentic motives.

rather than any *single* virtue that intuitively seems consistent with the motives that best enable agents to act authentically in particular circumstances.

In this respect, the indirection/self-effacement distinction put forth by Martinez looks as if it relies on an equivocation in granularity. Modern ethical theories like consequentialism are presumed to require the self-effacement of reasons grounded in abstract calculations, whereas virtue ethics is presumed to have no similar need to suppress reflection on its reasons because the reasons put forward for discussion are only those that are consistent with the motives that ought to guide agents in specific circumstances. Rather than a contrast between theories, the heavy lifting in the argument relies on shifting from *all-things-considered* justificatory reasons to reasons narrowly chosen to be consistent with motives we take to be authentic in particular examples. Without this shift in granularity, modern ethical theories can similarly put forward their own narrowly chosen reasons like promoting the good of friendship or treating others as ends in themselves, and then the argument advanced by Martinez only succeeds if it provides further explanation of why the theoretical justifications of virtue ethics are immune to risks of alienation compared to rival ethical theories. The possibility of providing such an explanation leads to one more intriguing reply to Keller's challenge.

4 Pettigrove's Value Pluralism

Of the attempts to refute Keller's claim that virtue ethics is self-effacing, the last of the three I will discuss comes from Glen Pettigrove.[32] The solution Pettigrove offers to rescue virtue ethics from self-effacement shares some prominent features with those presented by both Clark and Martinez, but it includes an additional element that deserves attention. These shared features include an emphasis on practical wisdom and the claim that a criterion of right action for virtue ethics (e.g. an action is right iff it would characteristically be performed by the virtuous person in relevant circumstances) can be interpreted by virtuous agents so that it is *transparent* in the sense that it allows its agents to see through the criterion to the morally salient details of their particular circumstances. Pettigrove notes that justifications for actions do not always attempt to identify what ought to serve as an agent's motive when she performs virtuous acts; thus, he argues that normative guidance that refers to what is virtuous will, "for a well-informed agent, simply direct her attention to a familiar constellation of value-constituting factors that make up the target of the virtue."[33] In effect, what Pettigrove's view shares with others is an attempt to illustrate how virtue ethics enables agents to pursue the acts recommended by its foundational justifications understood in *de re* terms rather than the *de dicto* terms that generate alienation because they abstract from the details of agents' lived experiences.

---

[32] Glen Pettigrove, "Is Virtue Ethics Self-Effacing?" *The Journal of Ethics* vol. 15, no. 3 (2011): 191-207.

[33] Ibid., p. 201.

The additional element in Pettigrove's proposal that merits attention is his emphasis on value pluralism. Though he is careful to acknowledge that the normative reasons provided by an ethical theory need not always function as motives for its agents, Pettigrove argues that the *fundamental* values of a theory must be able to serve as its major motives. This proviso creates a problem for ethical theories seeking to systematize the complexity of agents' lives, he argues, because:

> Someone who is a monist about goodness or value, who thinks that all apparently different types of value are really instances of or grounded in a single, underlying quality will have difficulty avoiding the charge of self-effacement because different action-guiding thoughts are required for different kinds of excellent action. The thoughts and attitudes one has when manifesting a type of goodness like creative self-expression will be different from those in which one is benevolently promoting another's good. […] So a virtue ethical account that would avoid self-effacement will need to allow for a plurality of values.[34]

Consider Stocker's hospital example. If Smith visits you because he cares about you as a friend, his otherwise authentic motive is arguably compromised if it turns out to depend upon a more fundamental value lurking in the background of his moral psychology. Even a version of virtue ethics like the eudaimonism defended by Annas will fail if this is true, according to Pettigrove, because the theory is ultimately governed by one fundamental value that must be self-effaced to avoid alienation in the motives of agents responding to goods that are related to particular virtues. Value pluralism, by contrast, allows agents to act for the sake of goods like friendship without a comprehensive value operating in the background to systematize agents' competing moral commitments. With each ethical value left to exert *direct normative influence* on agents without being mediated by a unified value, value pluralism looks like it can avoid the prospect of alienation without requiring that agents internalize a systematization of values through the acquisition of practical wisdom. Instead, Pettigrove's value pluralism avoids the threat of self-effacement because it proposes no systematization at a fundamental level.

This extra contribution from Pettigrove is important to acknowledge because it captures the sometimes misunderstood intent of Stocker's original argument against "modern" forms of systematic ethical theory.[35] Nevertheless, an appeal to value pluralism is not sufficient to save virtue ethics from self-effacement. Even a version of the (anti)theory that supplies agents with reasons and motives that refer directly to a

---

[34] Ibid., p. 200. As Pettigrove notes, Swanton (1997) also provides an argument for the role of value pluralism in avoiding self-effacement.

[35] As I argue in (Woodcock 2010), Stocker's prescription for avoiding moral schizophrenia is to approach what he considers to be a realm of plural and conflicting values without any kind of unified theoretical perspective that would, he claims, fail to directly capture the values in the manner that they deserve. Stocker outlines this view in: "Values and Purposes: The Limits of Teleology and the Ends of Friendship", *The Journal of Philosophy* vol. 78, no. 12 (1981), pp. 747–65; "Friendship and Duty: Some Difficult Relations," in *Identity, Character and Morality: Essays in Moral Psychology*, ed. Owen Flanagan and Amelie Oksenberg Rorty (Cambridge, Mass.: MIT Press, 1990), p. 219 – 234; and *Plural and Conflicting Values* (Oxford: Oxford University Press, 1992).

plurality of incommensurable values will, in at least some cases, require that agents suppress some of the most fundamental justifications for their actions. To see that this is so, consider the source of the alienation that creates a need for self-efface-ment. The problem for an agent like Smith occurs when you discover that some abstract, impersonal justification has played a role in motivating him to visit you. For example, if Smith is a consequentialist the problem arises because, even if he cares deeply for you, his visit is partly motivated by having reflected and confirmed that this action is compatible with a more general agenda to maximize the good. Similarly, for a Kantian agent, Smith may care for you as a friend for your own sake, but the motive for his visit will be partly constituted by having ensured that this act is compatible with broader ethical duties prescribed by the categorical imperative. It is in this respect that the key feature of Stocker's initial challenge is related to Wil-liams' claim that impartial ethical theories require "one thought too many" because they expect agents to reflect and justify actions performed for the sake of loved ones against the broader obligations of their foundational values.[36]

Yet with the source of this alienation made explicit, it becomes apparent that the root of the problem is not specific to modern or impartial theories. Even if Smith visits you because he cares passionately about *the particular value of friendship*, it will be a shock if he discloses that his motive for visiting you was based on careful reflection about how to most faithfully exhibit love for those he cares about and the result of this reflection happened to be visiting you in the hospital. His fundamental values could be an idiosyncratic mix of distinctly partial normative commitments, but if he reflects to seek justification from the overall scope of his values before acting, then Smith is still vulnerable to the objection that his motives are compro-mised when he chooses to visit you. The type of alienation that triggers a need for self-effacement is therefore incredibly difficult to avoid. Consider what it takes to avoid it: Smith would need to be an agent who performs actions like visiting you in the hospital *without reflecting on whether his actions are justifiable in the sense of exhibiting fidelity to the fusion of incommensurable values that he accepts as nor-matively significant*. This is not a plausible option. Whether an agent derives her justifications from a single, comprehensive value or an incommensurable set of dis-tinct values, she is still required to reflect and consider whether her actions are mor-ally defensible. To think otherwise is to confuse value pluralism with an implausible straw man – a view that would end up permitting agents to act for the sake of some particular value, like those related to the virtue of courage for example, without hav-ing to confirm that this act is compatible with the broader mix of values, incom-mensurable or not, to which the agent is committed. Smith might visit you without hesitation today, but tomorrow he might choose to honour the values associated with honesty and tell everyone he knows that you were in the hospital recovering from hemorrhoid surgery. The story need not even be true; perhaps he has decided to hon-our values associated with wit and tells everyone the hemorrhoid story despite the fact that you had your appendix removed. These scenarios are absurd only because

---

[36] Bernard Williams, "Persons, Character and Morality," in *Moral Luck* (Cambridge: Cambridge Univer-sity Press, 1981), 1-19, p. 17-18.

we *implicitly assume* that friends are bound by a complex web of interrelated values connected to privacy, justice and integrity. Proceeding as if agents can avoid reflecting on the values in this web and focus only on the most immediate details of their circumstances is not a viable strategy to seek refuge from self-effacement. It would be to accept an extreme form anti-theory that is beyond value pluralism. It would be a view that gives away far too much in terms of the fidelity that agents ought to exhibit to their overall normative commitments.[37]

Thus, value pluralism initially looks like it escapes from the threat of alienation because it has no definitive set of values to condemn if agents experience alienation. However, a mix of virtue ethics and value pluralism cannot have it both ways. Either it avoids the risk of alienation by conceding that agents need not reflect on the overall balance of values that determine the virtues, leaving the view without a credible way to give normative guidance, or it avoids being no more prescriptive than moment-to-moment intuitionism, leaving virtue ethics in the same position as other theories that require agents to reflect and ensure that they exhibit fidelity to their fundamental values (whether the values are part of a systematic theory or not).[38] Neither option is appealing to Pettigrove, but the latter is surely preferable for virtue ethics to remain normatively significant. Thus, even a version of virtue ethics faithful to Stocker's original vision remains vulnerable to the threat of self-effacement, the difference being only that it sometimes requires agents to separate their motives from justifications grounded in a complex mixture of incommensurable values rather than justifications from a single, foundational value. The result is the same: to honour their overall values in particular circumstances, agents must sometimes avoid being motivated by justifications that explicitly refer to these values. It is in this respect that virtue ethics must engage in the kind of reflection that invites self-effacement in order to remain normatively significant.

5 Conclusion

My aim has been to establish that virtue ethics is not immune to the difficulties associated with self-effacement compared to rival ethical theories. Nevertheless, I hope it is clear that recent attempts to rescue virtue ethics from self-effacement lead to key insights for an understanding of how agents ought to incorporate criteria of right action into their lived experience. A careful development of practical wisdom

---

[37] For comparison, it is helpful to consider the many different ways in which one might combine virtue ethics and anti-theory, few of which would be so radical as to reject some form of reflection on whether one's choices are in harmony with one's fundamental values. See: Robert B. Louden, "Virtue ethics and anti-theory," *Philosophia* vol. 20, no. 1 (1990): 93-114. For example, one might even endorse a pluralist version of virtue ethics that rejects the need for a criterion of right action, as has been proposed by John Hacker-Wright, and still retain a thoughtful type of overall critical reflection that invites self-effacement. See: "Virtue Ethics without Right Action: Anscombe, Foot, and Contemporary Virtue Ethics," *The Journal of Value Inquiry* vol. 44, no. 2 (2010): 209-224.

[38] Rebecca Stangl identifies a separate, though related, dilemma for particularist virtue ethics in, "A Dilemma for Particularist Virtue Ethics" *The Philosophical Quarterly* vol. 58, no. 233 (2007): 665-678.

is essential for agents to gain sensitivity to salient details in particular circumstances without having to constantly reflect on the theoretical justifications for their decisions. Additionally, it is important to consider the level of abstraction at stake when it comes to deciding whether the theoretical justifications for one's decisions are in conflict with the authenticity of particular motives (and, therefore, whether one ought to avoid reflecting on these justifications too frequently.) And, finally, the reason self-effacement is so hard to avoid is that if an ethical theory is to remain capable of offering normative guidance to agents, it must sometimes risk alienation in their particular commitments by prompting them to justify their choices against their most fundamental values, whatever the (singular or plural) fundamental value(s) might be that shape their chosen ethical theory.

With these insights in mind, what are we to make of the result that virtue ethics is not immune to self-effacement? There are various options. First, one could consider the result so improbable that it forces us to re-evaluate our understanding of self-effacement. It just cannot be true, on this option, that ethical theories as varied as Kantian ethics and value pluralism are vulnerable to self-effacement, so at some point in the literature a mistake in our conception of self-effacement must have been made. Second, one might provisionally accept the result but recommend that advocates of virtue ethics continue to search for a formulation of the theory that manages to avoid self-effacement. It is still possible that some version of virtue ethics not yet articulated can accomplish this aim without giving up on the kind of critical reflection that makes ethical theories normatively significant. Third, one might accept the result that virtue ethics is self-effacing yet claim that as a matter of contingent, empirical fact the conditions in which self-effacement is required by virtue ethics are much less common and/or onerous than those in which it is required by contemporary rivals to virtue ethics. In the wake of Stocker's original framing of the debate, advocates of virtue ethics have sought to establish that modern ethical theories are, as a matter of principle, uniquely vulnerable to self-effacement in a way that virtue ethics avoids entirely. However, this strategy need not be maintained. Those who defend virtue ethics could instead dig into the comparative details of implementing their own theory and those of their rivals in order to argue that the levels of self-effacement required by virtue ethics turn out to be substantially less burdensome.

Despite the availability of these options, I will close by suggesting a further alternative that I think advocates of virtue ethics ought to consider. This is to accept that self-effacement is a tolerable psychological challenge and need not be presumed to be detrimental to theories that require some compartmentalization between global justifications for right action and the motives that secure the authenticity of our particular commitments. This may seem a radical step given the trajectory of the literature, yet it is not obvious why advocates of virtue ethics remain committed to the claim that self-effacement is such an objectionable feature for an ethical theory. Given the influence of Stocker's original argument it is easy to presume that agents should not be directed to avoid reflecting on the justifications for their actions – that requiring any psychological indirection must be deeply injurious. Yet this is an unsupported assumption in the absence of empirical evidence, and it is one that risks begging important questions about moral psychology by starting from such a strong commitment to simplicity in our understanding of how minds of ideal agents ought

to function.[39] Consequently, it would not be unreasonable for an advocate of virtue ethics to remain provisionally agnostic about the harmfulness of self-effacement until we know more about just how taxing the challenge of self-effacement is for human agents.

While it is beyond the scope of this paper to defend this fourth option in detail, let me offer three reasons for thinking it is a sensible choice to consider. First, as we have seen in our discussion of Pettigrove, the conditions that trigger self-effacement are remarkably *ordinary* in the sense that any theory requiring agents to reflect on their overall normative commitments will sometimes need to insulate the motives of agents from constant critical reflection in order to preserve the authenticity of these motives. If the conditions that trigger self-effacement are this ordinary then it gives us some reason, I think, to start from the provisional assumption that the psychological processes involved are not pathological. Second, an expectation of simplicity in our moral psychology seems at odds with the fact that *general* human psychology is littered with many strange features in our ordinary patterns of thinking.[40] Given these precedents in non-moral contexts, it is unclear why we ought to expect simplicity in the moral psychology of agents put forward as archetypes of ethical theories. Finally, it seems reasonable to begin with a neutral burden of proof when making hypothesis about what is psychologically harmful and what is not. It is, of course, possible that self-effacement is psychologically harmful even if we do not currently have empirical evidence proving this to be the case. In the context of choosing between rival ethical theories, however, it seems prudent to begin with the view that contested features of our moral psychology are not so challenging that they are harmful until convincing evidence suggests otherwise.

All of this is to suggest that advocates of virtue ethics can allow themselves the luxury of remaining agnostic about whether it would be a problematic result to accept that their ethical theory is self-effacing.[41] The inertia of the literature may be

---

[39] For some this presumption is supported by Aristotle's claim that virtues are only exhibited if agents perform acts "knowingly" and "for themselves". (Book 2, section 4; Opt. cit. p. 115.) It is certainly clear from this section of the *Nicomachean Ethics* that Aristotle intends to exclude the possibility of agents being called virtuous if they perform virtuous acts without appropriate motives and stable character traits. Yet the extent to which this exclusion entails that virtuous agents must exhibit a *simple* moral psychology for their actions to be purposeful and deliberate is an open question – one that is beyond the scope of this paper.

[40] I am not referring here to the many cognitive biases to which we are susceptible, since one might plausibly claim that they should not be included within the parameters of virtuous human activity. I am referring to the curious features of human psychology that are likely part of adaptive patterns of reasoning, e.g. inattentional blindness, the paradox of choice, retrieval-induced forgetting, etc.. Perhaps the best example of a common pattern related to calculative vulnerability is the way we tend to generate novel insights in the shower when we rely on the brain's "default mode network" instead of concentrating directly on generating results. See: Arien Mack and Irvin Rock, *Inattentional Blindness* (Cambridge, MA: MIT Press, 2000); Barry Schwartz, *The Paradox of Choice: Why More is Less* (New York: Ecco, 2004); Patricia Román et al., "Retrieval-Induced Forgetting and Executive Control," *Psychological Science*, vol. 20, no. 9 (2009): 1053-1058; and Jessica R. Andrews-Hanna et al., "Evidence for the Default Network's Role in Spontaneous Cognition," *Journal of Neurophysiology* vol. 104, no. 1 (2010): 322-335.

[41] An alternate title for this paper might be, "How I Learned to Stop Worrying and Love Virtuous Self-Effacement." Yet this would be to overstate my case, which is only that advocates of virtue ethics consider remaining agnostic about whether self-effacement is a tolerable psychological challenge. Still, my thesis contrasts sharply with Damien Cox's claim that agent-based theories of right action are deficient

difficult to reverse, but advocates of virtue ethics are under no obligation to retain committed to the claim that they must avoid indirect features in our moral psychology at all costs. They can instead consider the possibility that self-effacement is an ordinary part of implementing ethical theory in social circumstances that require balancing critical reflection with authentic motives for action.

Finally, let me emphasize that nothing I suggest in this paper is incompatible with virtue ethics being a superior ethical theory compared to its rivals. The conclusion here is that there is no useful distinction to be drawn that gives virtue ethics a principled advantage in the context of indirect moral psychology, but this conclusion is consistent with other theoretical advantages the theory might possess. The insights virtue ethics offers when it comes to practical wisdom and the psychology of agents acting from authentic motives are considerable, so it is tempting to assume that these insights must give virtue ethics some distinct advantage. However, since other ethical theories (if they are sensible) can co-opt these insights, my recommendation is that ongoing debates between modern and virtue-based theories move on to more productive territory. Virtue ethics may well prevail in these debates by displaying theoretical advantages that give us reason to prefer it over modern theories. The argument I offer here aims only to make the case that these advantages are not obviously related to issues of self-effacement.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Footnote 41 (continued)

if they fail to exhibit deliberative transparency. See: "Agent-based Theories of Right Action," *Ethical Theory and Moral Practice* vol. 9, no. 5 (2006): 505-515.