

# Introduction

*Tyler Paytas and Tim Henning*

In the preface to *On What Matters*, Derek Parfit refers to Sidgwick and Kant as his two masters—the historical figures who had the greatest influence on his ethical views (2011, xxxiii). This reverence is hardly insignificant given that Parfit is among the most brilliant philosophical minds of the past 100 years and thus highly influential in his own right. Parfit claims that Kant is the greatest moral philosopher since the ancient Greeks, adding the memorable remark that “in the cascading fireworks of a mere forty pages, Kant gives us more new and fruitful ideas than all the philosophers of several centuries” (2011, xxxiii, 183). Equally unreserved in praise of his other master, Parfit boldly claims that Sidgwick’s *Methods of Ethics* is “the best book on ethics ever written” (2011, xxxiii).

While some may classify these remarks as overstatements, they are certainly not unfounded. Kant’s central importance in the history of ethics is undeniable; his rationalist, deontological approach to moral philosophy has spawned generations of disciples who have forcefully interpreted and championed Kantian doctrine (Rawls 1971; O’Neill 1975; Herman 1993; Baron 1995; Wood 1999). Further, Kant’s writings have inspired a distinctive and influential brand of metaethical constructivism (Korsgaard 1996; Reath 2006; Hill 2012), while also making a significant impact in applied areas such as medical ethics (Cohen 1986), animal ethics (Regan 1983), and law (Shiffrin 2014).

Although Sidgwick’s affect on moral philosophy has not been as pervasive or consistent as Kant’s, his work was a strong influence on some of the most important ethical theorists of the twentieth century, including Moore (1903), Ross (1930), Rawls (1971), and Parfit (1984; 2011). Indeed, the Sidgwickian ideas permeating Parfit’s writings are partly responsible for a recent surge of interest in Sidgwick. Since the publication of *On What Matters* in 2011, several monographs on Sidgwick have been published by major presses (Phillips 2011; Nakano-Okuno 2011; de Lazari-Radek and Singer 2014; Crisp 2015), with others in the pipeline (Kagan ms; Skelton forthcoming). This is in addition to important volumes of broader scope that nonetheless contain extensive coverage of

## 2 Tyler Paytas and Tim Henning

Sidgwick (Hurka 2014; Schultz 2017). The result of these developments has been an expansion of the paradigm for study of classical utilitarianism, from the duo of Bentham and Mill to the triumvirate of Bentham, Mill, and Sidgwick.

Still, Parfit's dual exaltation of Kant and Sidgwick may initially seem surprising given that the two philosophers champion ethical theories commonly seen as antithetical to each other (deontology and consequentialism). But this begins to look less puzzling in light of Parfit's deeply felt desire to find convergence within moral philosophy. Like Sidgwick, Parfit was highly sensitive to the skeptical threat posed by genuine philosophical disagreement, which seems especially pervasive in ethics. If the most careful and reflective thinkers are in a state of perpetual disagreement over fundamental ethical principles, this casts doubt on our ability to grasp objective ethical facts and even whether such facts exist at all. Hence, philosophers such as Parfit, who believe in objective and knowable ethical truths, incur a philosophical burden of explaining why the apparent disagreement does not warrant skepticism. One strategy for avoiding this conclusion is to carefully study the arguments and insights of figures such as Kant and Sidgwick in order to better understand the respective theories and search for points of contact between them. The present volume is motivated in large part by this idea.

At first blush, it might appear that comparative study of Kant and Sidgwick will only serve to deepen worries about irresolvable ethical disagreement. After all, Kant is not ambiguous regarding his disdain for teleological and consequentialist methods: "Woe to him who creeps through the serpent-windings of utilitarianism" (Kant 1996, MM 6:332). Though less vitriolic, Sidgwick is equally doubtful of deontological approaches, describing them as "to some extent doubtful and confused; and sometimes, even when clear, as merely dogmatic, unreasoned, and incoherent" (ME xvii). Moreover, Sidgwick offers sharp criticisms of some of Kant's central arguments and views. For instance, in considering Kantian moral theology as a possible solution to the apparent rational conflict between prudence and morality, Sidgwick writes:

I cannot fall back on the resource of thinking myself under a moral necessity to regard all my duties *as if they were* commandments of God, although not entitled to hold speculatively that any such Supreme Being really exists. I am so far from feeling bound to believe for purposes of practice what I see no ground for holding as a speculative truth, that I cannot even conceive the state of mind which these words seem to describe, except as a momentary half-willful irrationality, committed in a violent access of philosophic despair.

(ME 507 n. 3)

Other criticisms include the charge that Kant equivocates in his appeals to freedom as the cornerstone of morality (ME 511–516), as well as the allegation of a “paralogism” in Kant’s attempt to derive a principle of benevolence from the idea of humanity as an end in itself (ME 389–90).

Not all of Sidgwick’s references to Kant were critical, however. After presenting his ethical axioms in Book III of the *Methods*, Sidgwick notes that his confidence in them is buoyed by the agreement of those moralists who have been “most in earnest” in their philosophical investigations. Along with Samuel Clarke, Sidgwick cites agreement from Kant, who he describes as “especially noted for his rigour in separating the purely rational element of the moral code” (ME 384–85). While acknowledging some disagreement in the details, Sidgwick views Kant as an important source of corroboration for his principle of Rational Benevolence as well as his axiom of Justice, which he takes to be a corollary of the Universal Law formulation of Kant’s Categorical Imperative (ME 386). Sidgwick also expresses a feeling of kinship with Kant regarding the philosophical desire for system and coherence and the resulting difficulties: “Oh how I sympathise with Kant! With his passionate yearning for synthesis and condemned by his reason to criticism” (Sidgwick and Sidgwick 1906, 177).

Although Kant was hostile toward utilitarianism, there is reason to think that he would have appreciated several aspects of Sidgwickian ethics (as compared to the systems put forth by Bentham and Mill). One respect in which Sidgwick differs from his utilitarian predecessors is that he recognizes the need for a rational intuition that can provide a secure foundation for a principle of universal benevolence. Appeals to emotion or the heroic nature of self-sacrifice were not enough for Sidgwick. He memorably expresses this concern in an autobiographical note from the preface to the sixth edition of the *Methods*:

It was no use to say that if I was a moral hero I should have formed a habit of willing actions beneficial to others which would remain in force, even with my own pleasure in the other scale. I knew that at any rate I was not the kind of moral hero who does this without reason; from *blind* habit. Nor did I even wish to be that kind of hero; for it seemed to me that that kind of hero, however admirable, was certainly not a philosopher. I must somehow *see* that it was right for me to sacrifice my happiness for the good of the whole of which I am a part.

(ME xviii)

The preference for rationalism over sentimentalism is one of the factors that seemingly brings Sidgwick closer to Kant than initial appearances suggest. Another point of contact between the two is their recognition of

#### 4 Tyler Paytas and Tim Henning

the philosophical significance of conflicts between duty and self-interest. Kant and Sidgwick are generally in agreement that human happiness consists in agreeable states of consciousness, and that there is no reliable connection (at least in the earthly realm) between adhering to duty and partaking in such happiness. That said, the two philosophers draw divergent conclusions from the apparent inevitability of moral-prudential conflicts. Whereas Kant believes that the rational authority of morality could not be undermined by conflicts with self-interest, Sidgwick sees an unresolved “dualism of the practical reason” that significantly threatens the enterprise of philosophical ethics (more on this below).

In an essay on Kant’s *Critical Philosophy*, Sidgwick confesses that “I have often longed to call up the philosopher of Königsberg himself,” to ask about interpretations of his views (1883, 71). Although this was an idle wish, we may perhaps obtain some of the fruits of this imagined encounter by exploring the work of both philosophers together. Such comparative historical studies are useful not merely as a means of seeking philosophical convergence but also for clarifying and strengthening our own thinking on the relevant topics. As C. D. Broad explains in the introduction to his classic work *Five Types of Ethical Theory*:

The best preparation for original work on any philosophic problem is to study the solutions which have been proposed for it by men of genius whose views differ from each other as much as possible. The clash of their opinions may strike a light which will enable us to avoid the mistakes into which they have fallen; and by noticing the strong and weak points of each theory we may discover the direction in which further progress can be made.

(Broad 1930, 1–2)

In this spirit, the present volume aims to advance both contemporary ethical thought and our understanding of Kant and Sidgwick by using the work of each philosopher to illuminate that of the other. In addition to the question of the extent to which common ground between deontological and consequentialist approaches is possible, the juxtaposition of Kant and Sidgwick will shed light on several important topics including free will, moral epistemology, and moral theology.

The volume begins in Part I with discussions of Kant’s and Sidgwick’s views on normative, practical ethics, and of possible connections and divergences between them.

Anthony Skelton’s chapter offers an overview of Kant’s and Sidgwick’s views on central issues in practical ethics. Skelton focuses especially on their contrasting positions on the morality of lying and the demandingness of the duty of beneficence. In addition, he investigates how far the two authors can claim to be in general agreement with the morality of common sense. This last point is of major importance because both Kant

and Sidgwick hold metaethical views that lead one to expect that divergence from common sense should be a rare exception rather than the norm. According to Skelton, it turns out that Kant and Sidgwick do disagree with each other and with common sense on a number of counts. But all in all, he argues that Sidgwick's approach is more promising when it comes to accommodating common-sense verdicts.

Tim Henning's contribution highlights a point of agreement between Sidgwick and Parfit, on the one hand, and Kant on the other. Among many other things, Sidgwick has become famous for his differentiated views on the question of when a criterion of rightness should serve as the motive for moral action. In particular, Sidgwick argues that the true criterion of rightness may in fact demand that some other consideration should be motivating rather than the criterion itself. Henning follows Parfit's careful analysis of the ways in which such practical self-defeat can occur, and he suggests that Kant's Categorical Imperative may be subject to a particular form of indirect self-defeat, much like Sidgwick and Parfit have claimed with regard to utilitarianism. Henning argues that this self-defeat plays a major role in the basic distinctions and the overall architecture of Kant's *Doctrine of Right* (*Rechtslehre*). Henning ends with the suggestion that this late Kantian work should be read as developing a form of 'indirect deontology' on the model of indirect consequentialism.

Part II of the volume follows up with discussions of the topics of practical knowledge, rational deliberation, and moral motivation.

Carla Bagnoli begins her chapter by emphasizing similarities in Sidgwick's and Kant's theoretical projects. As Bagnoli stresses, both authors reject a naturalistic outlook, according to which moral demands can be fully explained in terms of the nature of humans and their circumstances. Likewise, both Kant and Sidgwick reject certain forms of intuitionism. Against naturalistic and (some) intuitionist views, they insist on the practical force of moral judgments, as directly capable of binding the rational subject's will. At the same time, however, both philosophers reject the idea that moral judgments are simply expressions of antecedently existing motivating states of the subjects, thus making moral judgments depend on contingent psychologies. After Bagnoli carefully lays out how Sidgwick's version of intuitionism is supposed to live up to these requirements, she suggests that Kant's view ultimately holds greater promise of success. This, Bagnoli argues, is because Kant conceives of moral knowledge as a kind of practical knowledge, i.e., as a knowledge of principles that arises directly out of the nature of practical decision making.

David Phillips's contribution explores an important line of argument in Sidgwick's metaethics. Sidgwick endorses the non-naturalist view that terms like 'ought' cannot be paraphrased in naturalistic terms (e.g., in terms of desire-fulfillment), and he also holds the Kantian view that reason alone can motivate. Now, these two views are presented by Sidgwick in a way that suggests that there is an argument linking them. Phillips

discusses whether there is such a connecting argument and, if so, what this argument is. He begins by arguing, contra J. Schneewind (1977), that there is indeed an argument linking the non-naturalness of ‘ought’ and the Kantian (or anti-Humean) view of moral motivation. Phillips then argues against one reconstruction (proposed by, among others, R. Shaver (2000)) of this argument. According to this reading, Humeans need to deny non-naturalness. In order to explain away apparent conflicts between reason and passion, they need to re-interpret the verdicts of reason as being concerned with prosaic facts about desires after all. Phillips rejects this reconstruction and offers a different interpretation. On Phillips’s reading, it is Kantian accounts of motivation that are committed to non-naturalness. This is because the Kantian view of moral motivation receives its intuitive support from the appearance that ‘ought’-judgments play a distinctive role in motivation. And this appearance could not be vindicated if such judgments could be paraphrased away by statements expressing or reporting ordinary, Humean motivation. After reconstructing the argument, Phillips then shows that Sidgwick’s positive account of moral motivation very closely resembles that of a contemporary non-naturalist, T. M. Scanlon (1998; 2014).

Part III of the volume offers detailed discussions of Sidgwick’s and Kant’s relation to an intuitionist moral epistemology, and of the significance of *peer disagreement* in moral matters—including apparent disagreements between Kant and Sidgwick.

Roger Crisp’s contribution discusses the question of whether there is a fundamental disagreement between Sidgwick’s and Kant’s moral epistemologies. First, Crisp discusses the character of Sidgwick’s philosophical intuitionism, contrasting it with perceptual and dogmatic intuitionism. Famously, Sidgwick himself considered Kant an ally in this epistemological approach—a claim that has been met with much criticism. Crisp, however, replies to these criticisms and argues that considering Kant as a philosophical intuitionist may be tenable after all.

One apparent difference between Sidgwick and Kant (which is also discussed in Crisp’s chapter) concerns the significance of *dissensus* or disagreement. Sidgwick is often cited as an adherent of a conciliatory view regarding peer disagreement. Specifically, he is thought to have held that in cases of peer disagreement, suspension of judgment is called for. Robert Shaver carefully reconstructs Sidgwick’s views, and he proposes an interpretation according to which peerhood is not the default assumption. While Sidgwick includes among one’s epistemic peers all those for whom one has no more reason to suspect of error than oneself, Shaver proposes a view according to which, in a case of disagreement, peerhood has to be established by positive argument. Shaver then goes on to argue that this modified interpretation fits well with Sidgwick’s discussion of intertemporal disagreement in the history of thought. Importantly,

Shaver concludes that Sidgwick need not have regarded Kant as his epistemic peer in the domain of philosophical ethics.

The topic of Part IV is the ethical significance of free will. Whereas Kant famously views freedom as central to morality, Sidgwick argues that the question of free will is ultimately insignificant for ethics. In Sidgwick's view, when faced with a choice between alternative modes of conduct, the question of whether all of our decisions and actions are ultimately determined by antecedent causal forces is irrelevant. Sidgwick also accuses Kant of equivocating between two distinct notions of freedom—one which is manifest whenever we are guided by reason rather than impulse, and another which is exhibited when we exercise the capacity to choose between right and wrong regardless of which one we choose (ME 511–512).

In her chapter, Mariko Nakano-Okuno argues that while Sidgwick's accusation of equivocation was off base, his skepticism about the practical significance of free will was largely on target. Nakano-Okuno analyzes and defends Sidgwick's argument for the claim that our capacity to act from rational principles, or in Kantian terms, to set universal laws for ourselves, does not require the assumption of our will's complete freedom from antecedent biological and environmental causes. Further, while moral emotions such as guilt and remorse may lose their motivational force in the absence of belief in freedom, we can still be motivated to act rightly by the powers of imagination, sympathy, and love of goodness. Regarding the issue of holding others responsible, Nakano-Okuno defends the Sidgwickian idea that prospectively oriented practices of censure and punishment, which do not require the assumption of free will, can be every bit as effective as retributive concepts and practices that do presuppose free will. Nakano-Okuno optimistically concludes that if Kant were to jettison his commitment to the moral significance of freedom, then there would be no insurmountable discrepancy between his ethical system and that of Sidgwick.

In his contribution, Paul Guyer agrees that freedom is the fundamental source of divergence between Kant and Sidgwick. Indeed, Guyer suggests that the disagreement between Kant and Sidgwick over the moral significance of freedom is central to the general conflict between utilitarianism and deontological ethics, which he sees as largely concerning the derivation and scope of the duty of benevolence. However, Guyer does not see the downplaying of freedom as a live option for Kantian ethics. Guyer's central claim is that Sidgwick fails to appreciate the merits of Kant's having derived the duty of benevolence not from the value of happiness but rather from the value of freedom. This is a crucial point insofar as the freedom-based conception of morality entails constraints on the promotion of happiness that constitute the core tension between deontological and utilitarian approaches.

The promotion of happiness is also pertinent to the final part of the volume (Part V), which addresses Kant's and Sidgwick's respective views on ultimate ends. As a consequentialist, Sidgwick conceives of ethics as largely a matter of determining the ultimate end of reasonable conduct. And while Kant believes that morality is composed of universal rational principles that apply irrespective of consequences, he acknowledges that human practical reason has a deep need to posit an ultimate rational end (1998, Rel 6:5). Both philosophers recognize a difficulty arising from the idea of an ultimate rational end, which is that such an end would seemingly need to contain two elements that are not reliably connected—duty (which for Sidgwick is a matter of impartial benevolence) and one's own happiness. The fact that dutiful conduct seems to at least occasionally require sacrifice of one's own happiness is perhaps the most fundamental problem for philosophical ethics (see Shaver 1999; White 2002; Crisp 2019).

One way of responding to this problem is to claim that the appearance of moral-prudential conflicts is illusory. This is the approach taken by ancient eudaimonist theories, which hold that the most morally virtuous life is also the happiest. However, Kant and Sidgwick are both highly skeptical of this approach. Among their chief objections are that eudaimonist views either posit implausible conceptions of happiness or they are objectionably egoistic. These critiques of eudaimonism are the subject of Sukaina Hirji's contribution. After outlining and analyzing Kant's and Sidgwick's objections, Hirji argues that there are resources available within the eudaimonist tradition to overcome them. Hirji focuses specifically on Aristotelian eudaimonism, which is the version she takes to be best equipped to meet the challenges pressed by Kant and Sidgwick. She advocates an interpretation of Aristotle involving a version of "weak eudaimonism," in which acting virtuously is a matter of performing right actions with the appropriate reasons and desires, and that acting in this manner is constitutive of human happiness. Hirji argues that this interpretation allows us to see why the Aristotelian agent need not be objectionably egoistic. As to whether the proposed conception of happiness is sufficiently plausible, Hirji provides some considerations in defense of Aristotle while acknowledging that the disagreement over the nature of human happiness is not easily resolvable.

One of the puzzling similarities between Kant and Sidgwick is that, despite their mutual advocacy of secular ethics (as opposed to divine command approaches), they both conclude that belief in God is morally necessary. In his chapter, Tyler Paytas argues that this appeal to God is related to their shared views about the idea of an ultimate rational end. As mentioned earlier, both philosophers recognize that one's own happiness must be included in any plausible conception of an ultimate rational end, but also that apparent moral-prudential conflicts raise serious difficulties on this score. Belief in a divine being who aligns duty and

self-interest with post-mortem rewards is one way out of this predicament. While Kant appeals to these considerations as practical grounds for faith, Sidgwick expresses skepticism. However, Paytas argues that Sidgwick's considered position is more similar to Kant's than initial appearances suggest. Paytas argues further that the predicament that leads Kant and Sidgwick to their need for faith can be avoided by abandoning their hedonistic accounts of well-being.

Kant's *Critique of Practical Reason* ends on a memorable note of inspiration: "Two things fill the mind with ever new and increasing admiration and reverence, the more often and more steadily one reflects on them: the starry heavens above me and the moral law within me" (1997, CpV 5:162). Sidgwick was less hopeful. At the conclusion of the first edition of the *Methods*, he somberly describes the implications of the dualism of practical reason: "The Cosmos of Duty is thus really reduced to a Chaos: and the prolonged effort of the human intellect to frame a perfect ideal of rational conduct is seen to have been foredoomed to inevitable failure" (ME1 473). In the preface to his *Volume Three*, Parfit reaffirms a qualified optimism that falls somewhere in between Kant and Sidgwick. Quoting his own earlier remarks concerning the skeptical challenge from ethical disagreement, Parfit writes: "Non-Religious ethics is at a very early stage. We cannot yet predict whether . . . we will all reach agreement. Since we cannot know how Ethics will develop, it is not irrational to have high hopes" (2017, xii; 1984, 454). Integrated study of Parfit's two masters may result in progress toward the convergence he hoped to find. But even if it does not, it will undoubtedly improve our understanding of the most important disagreements so that we might continue our ethical investigations with greater clarity and sympathy.

## References

- Baron, M. 1995. *Kantian Ethics Almost Without Apology*. Ithaca: Cornell University Press.
- Broad, C. D. 2013 [1930]. *Five Types of Ethical Theory*. New York: Routledge.
- Cohen, C. 1986. "The Case for the Use of Animals in Biomedical Research." *The New England Journal of Medicine* 314: 865–869.
- Crisp, R. 2015. *The Cosmos of Duty: Henry Sidgwick's Methods of Ethics*. Oxford: Oxford University Press.
- Crisp, R. 2019. *Sacrifice Regained: Morality and Self-Interest in British Moral Philosophy from Hobbes to Bentham*. Oxford: Oxford University Press.
- de Lazari-Radek, K., and P. Singer. 2014. *The Point of View of the Universe: Sidgwick and Contemporary Ethics*. Oxford: Oxford University Press.
- Herman, B. 1993. *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press.
- Hill, T. 2012. "Kantian Constructivism in Ethics." *Ethics* 99 (4): 752–770.
- Hurka, T. 2014. *British Ethical Theorists from Sidgwick to Ewing*. Oxford: Oxford University Press.

- Kagan, S. Unpublish ms. *Thinking Through Sidgwick*.
- Kant, I. 1996. *Metaphysics of Morals*. Edited by Mary Gregor. Cambridge: Cambridge University Press.
- Kant, I. 1997. *Critique of Practical Reason*. Edited by Mary Gregor. Cambridge: Cambridge University Press.
- Kant, I. 1998. *Religion Within the Boundaries of Mere Reason*. Edited and translated by A. Wood and G. Di Giovanni. Cambridge: Cambridge University Press.
- Korsgaard, C. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Moore, G. E. 2003 [1903]. *Principia Ethica*. Cambridge: Cambridge University Press.
- Nakano-Okuno, M. 2011. *Sidgwick and Contemporary Utilitarianism*. Basingstoke: Palgrave Macmillan.
- O'Neill, O. 1975. *Acting on Principle: An Essay on Kantian Ethics*. New York: Columbia University Press.
- Parfit, D. 1984. *Reasons and Persons*. Oxford: Clarendon Press.
- Parfit, D. 2011. *On What Matters*, 2 vols. Oxford: Oxford University Press.
- Parfit, D. 2017. *On What Matters*, vol. 3. Oxford: Oxford University Press.
- Phillips, D. 2011. *Sidgwickian Ethics*. Oxford: Oxford University Press.
- Rawls, J. 1999 [1971]. *A Theory of Justice*. Cambridge, MA: Belknap Press of Harvard University Press.
- Reath, A. 2006. *Agency and Autonomy in Kant's Moral Theory*. Oxford: Oxford University Press.
- Regan, T. 1983. *The Case for Animal Rights*. Berkeley: University of California Press.
- Ross, W. D. 1930. *The Right and the Good*. Oxford: Clarendon Press.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Scanlon, T. M. 2014. *Being Realistic About Reasons*. Oxford: Oxford University Press.
- Schneewind, J. B. 1977. *Sidgwick's Ethics and Victorian Moral Philosophy*. Oxford: Clarendon Press.
- Schultz, B. 2017. *The Happiness Philosophers: The Lives and Works of the Great Utilitarians*. Princeton: Princeton University Press.
- Shaver, R. 1999. *Rational Egoism: A Selective and Critical History*. Cambridge: Cambridge University Press.
- Shaver, R. 2000. "Sidgwick's Minimal Metaethics." *Utilitas* 12: 261–277.
- Shiffrin, S. 2014. *Speech Matters: Lying, Morality and the Law*. Princeton: Princeton University Press.
- Sidgwick, A., and E. M. Sidgwick. 1906. *Henry Sidgwick: A Memoir*. London: Macmillan.
- Sidgwick, H. 1874. *The Methods of Ethics*. 1st ed. London: Macmillan.
- Sidgwick, H. 1883. "A Criticism of the Critical Philosophy." *Mind* 8 (29): 69–91.
- Sidgwick, H. 1981 [1907]. *The Methods of Ethics*. 7th ed. Indianapolis: Hackett.
- Skelton, A. forthcoming. *Sidgwick's Ethics*. Cambridge: Cambridge University Press.
- White, N. 2002. *Individual and Conflict in Greek Ethics*. Oxford: Oxford University Press.
- Wood, A. 1999. *Kant's Ethical Thought*. Cambridge: Cambridge University Press.