

# Ethics as a Science

Diana Mertz Hsieh (diana@dianahsieh.com)

Ethics (Phil 5100, Bovens)

3 May 2003

In comparing the three major ethics of Western philosophy, i.e. those of Aristotle, Kant, and Mill, the most striking differences are found in their basic substantive theories. Aristotelian ethics strives for the rational flourishing of the individual within the polis. Kant's ethics aims for a will bound only by respect for universal law. Mill's utilitarianism seeks the collective happiness of humankind. These ultimate ends are central to each philosopher's moral theory, as they serve as the moral standard against which actions, principles, and characters are judged.

These abstract moral ideals, however, do not stand alone. Each philosopher also employs a model of moral judgment for translating the abstract standard into the concrete choices of everyday life. Aristotle requires a moral person to make virtue second nature through the cultivation of proper habits, skills, and feelings. Kant envisions a straightforward process of logical inference about the maxims of action. Mill appeals to intermediate generalizations based up comparisons of collective utilities. These three different accounts of moral judgment, in conjunction with their corresponding moral standard, largely determine each philosopher's basic vision of the structure of ethics and the nature of moral life.

Looking broadly, both Kant and Mill share an ambitious conception of ethics as the straightforward, algorithmic application of a universal rule to common moral questions. In sharp contrast, Aristotle explicitly denies that the complexity of ethics is amenable to either rigorous universality or simple inferential judgments. In other words, Kant and Mill seem to share the Enlightenment hope of transforming ethics into a science, while Aristotle happily regards it as an art. So we might ask: Can Kant and Mill live up to the hope of a science of ethics?

## **Moral Standard in Kant: The Categorical Imperative**

In the *Foundations of the Metaphysics of Morals*, Kant begins his substantive discussion of ethics by declaring that “Nothing in the world—indeed nothing even beyond the world—can possibly be conceived which could be called good without qualification except a *GOOD WILL*” (Kant 1985, 393). Kant's argument for a good will as the essence of morality is largely one of exclusion and logical dependence. He notes that generally beneficial talents like “intelligence, wit, and judgment,” qualities like “courage, resoluteness, and perseverance,” and gifts like “power, riches, honor, and even health” may become “extremely bad or harmful” without a good will (Kant 1985, 393). Even the “moderation in emotions and passions, self-control, and calm deliberation” advocated by ancient Greek philosophers will make a “villain... far more dangerous [and] abominable” (Kant 1985, 394). In fact, all such qualities are good only insofar as they “presuppose a good will”—and thus only a good will has “intrinsic unconditional worth” (Kant 1985, 394, 393). As an unconditional good, a will need not be effective, competent, or useful, because, as Paton explains, it is good “without limitation or qualification or restriction” (Kant 1985, 394; Paton 1965, 34). So even if completely incapable of achieving its ends, the good will would “sparkle like a jewel all by itself, as something that had its full worth in itself” (Kant 1985, 394).

Despite the temptation, we ought not adopt the common interpretation of this argument as claiming that “nothing that is good without qualification ever combines with anything to produce

a bad result” (Ross 1954, 10-11). After all, Kant is wholly undisturbed by the prospect that a good will acting from duty might cause terrible harms due to unfortunate outside circumstances. In his essay “On A Supposed Right To Tell Lies From Benevolent Motives,” Kant argues that even when faced with a deranged murderer banging on the door demanding to know whether his intended victim is inside, the moral course of action is to tell the truth (Kant 2003). Truthfulness is “a sacred unconditional demand of reason” which is “not to be limited by any expediency”—even when the “expediency” in question is the life of an innocent person (Kant 2003, P8). Thus clearly a good will cannot be unconditionally good because it always produces good consequences. Instead, Kant’s argument here better understood as claiming that while ordinary goods themselves become “the source of an additional badness” when directed by a bad will, the moral status of a good will cannot be so tainted (Paton 1965, 38). Kant’s basic observation here wields some intuitive force, as virtues such as efficiency and ingenuity seem to assume the stench of evil when used for evil ends such as genocide and torture.

However, even so interpreted, Kant’s argument is not entirely convincing for two basic reasons. First, the judgment that a good will is unconditionally good begs the question by presuming that consequences count for absolutely nothing in moral evaluation. Without this presumption against consequentialism, we could argue that the purity of a good will is corrupted to the extent that it disregards the likely consequences of action by, for example, truthfully revealing the location of intended victims to a murderer. In other words, the will that would choose (paraphrasing Hume) the destruction of the whole world to the slightest impurity could hardly be said to “sparkle like a jewel all by itself” (Kant 1985, 394). Second, Kant wrongly supposes that ethics requires an unconditional good at all. To avoid an infinite regress of ends, ethics merely needs an ultimate, intrinsically good end. A good that is always, no-matter-what good seems superfluous.<sup>1</sup>

Perhaps more problematically, Kant manages to rule out happiness as the final end in ethics simply based upon his ideas about the structure of the subject itself. In the opening pages of the *Foundations*, Kant divides ethics into “rational” and “empirical” parts based upon his vital metaphysical distinction between a priori and a posteriori knowledge (Kant 1985, 388). He then defines his purpose as the construction of a “pure moral philosophy which is completely freed from everything which may be only empirical” and thus wholly constituted by universal and necessary a priori truths (Kant 1985, 389). In fact, Kant claims that any concern for empirical matters “spoils the purity of morals” by mixing “pure principles with empirical ones” (Kant 1985, 390). The basic problem with this schema of ethics is that Kant never demonstrates either that his metaphysical distinction between the rational and the empirical applies ethics or that ethics contains a priori truths (Wilkinson 2001, P7).

Significantly, this mapping of ethics conveniently allows Kant to dismiss Aristotle, Mill, and all other philosophers concerned with human happiness as simply confused about the nature of ethics. Happiness, Kant argues is “so indefinite that, although each person wishes to attain it, he can never definitely and self-consistently state what it is that he really wishes and wills” (Kant 1985, 418). Such fuzziness is the product of the wholly empirical nature of happiness, such that a person “cannot... act according to definite principles so as to be happy, but only according to empirical counsels (e.g. those of diet, economy, courtesy, restraint, etc.) which are shown by experience to best promote well-being on the average” (Kant 1985, 418). Such contingent “counsels” do not meet Kant’s pure standard of “commands” of reason which “infallibly and

---

<sup>1</sup> However, Kant’s requirement of an unconditional good does make sense given his conception of ethics as grounded in universal, a priori, necessary truths, as discussed in his exclusion of happiness as a final end of ethics.

universally” determine the morality of action (Kant 1985, 418). But of course, neither Aristotle nor Mill share these meta-ethical standards. In this way, Wilkinson notes, “Kant, without much ado, dislodges happiness as the jewel in the crown of morality simply by drawing the boundaries of moral science in a particular way” (Wilkinson 2001, P6).

Leaving aside such worries for the moment, we can continue onward to the Categorical Imperative by noting that the argument for the good will as the only unconditional good is supposed to show that the moral worth of an action cannot depend upon any “hoped-for effect” or inclination (Kant 1985, 400). Rather, moral worth is wholly determined by the principle or maxim of the will that motivates action (Kant 1985, 399). More specifically, a good action is motivated by the formal, a priori “principle of volition as such,” namely “the necessity to do an action from respect for law” (Kant 1985, 400). Although this connection between the perfectly good will and acting from respect for law is tenuous at best, Kant’s conclusion is plain as day. Morality consists of the recognition of and obedience to the Categorical Imperative: “Act as though the maxim of your actions were by your will to become a universal law of nature” (Kant 1985, 421).

Later in the *Foundations*, Kant muddies the clarity of the Categorical Imperative by introducing two additional versions of the law. The three versions of the Categorical Imperative are thus as follows:

1. Principle of Universalizability: “Act as though the maxim of your actions were by your will to become a universal law of nature” (Kant 1985, 421).
2. Principle of Humanity: “Act so that you treat humanity, whether in your own person or that of another, always as an end and never as a means only” (Kant 1985, 429).
3. Principle of Kingdom of Ends: “Act as if your maxims should serve at the same time as universal law (for all rational beings)” (Kant 1985, 438).

Kant claims that these three principles are “fundamentally only so many formulas of the very same law” such that “each of them unites the others in itself” (Kant 1985, 436). Unfortunately, the equivalence of these three principles is far from self-evident—and Kant never offers any demonstration of the logical transformations of one to the other. (And as we shall see, the various versions of the Categorical Imperative can and do generate contradictory duties.)

Irrespective of the particular formula used, the basic function of the Categorical Imperative in Kant’s ethics is to illuminate the duties of all rational beings. Some of those duties are perfect (or impermissible), meaning that we are always obliged to conform our actions to them. Others are imperfect (or meritorious) duties, requiring only that we act upon them sometimes. Such duties are revealed in the process of moral judgment.

Considering the whole of Kant’s attempted justification of the Categorical Imperative, we see the same drive to transform ethics into a science—and the same appeal to a priori, universal truths—that characterizes Kant’s metaphysics. But while we might applaud his effort, we cannot congratulate him on his success. As we have noted throughout this discussion, his account of the foundations of ethics contains too many presumptions, glosses, mysteries, and logical leaps to justify the label “science.”

### **Moral Judgment in Kant: The Algorithm of Logical Inference**

In the Second Section of the *Foundations*, Kant applies the first two formulations of the Categorical Imperative to his four standard examples of suicide, false promises, developing talents, and helping others. In addition to establishing the relevant perfect and imperfect duties,

these case studies reveal Kant’s vision of moral judgment to be that of a simple and straightforward process of logical inference.

A brief review of the examples shows the four-step algorithm used with the Principle of Universalizability:

1. Suicide: A man “reduced to despair by a series of evils” wishes to commit suicide (Kant 1985, 422). However, when he attempts to universalize the maxim “For love of myself, I make it my principle to shorten my life when by a longer duration it threatens more evil than satisfaction,” he finds contradiction in the idea that the feeling “whose special office is to impel the improvement of life” could also cause of the destruction of life (Kant 1985, 422).
2. False Promises: A man is “forced by need” to borrow money, but knows that he will not be able to repay it (Kant 1985, 422). In order to secure a loan, he considers offering a false promise. When he attempts to universalize the maxim of his action, “When I believe myself to be in need of money, I will borrow money and promise to repay it, although I know I shall never be able to do so,” he finds contradiction in the idea itself, for such a universal law would undermine all trust in promises (Kant 1985, 422).
3. Talents: A man with a particularly useful talent finds that he would prefer lazy indulgence in pleasure to cultivating this talent. So could the maxim of “neglecting... gifts” for the sake of “idle amusement” become a coherent universal law? Although such “a system of nature could indeed exist,” Kant claims, the man cannot coherently *will* it because “as a rational being, he necessarily wills that all his faculties should be developed, inasmuch as they are given to him and serve him for all sorts of purposes” (Kant 1985, 423).
4. Helping: A well-off man who sees others struggling with “great hardships” denies any concern and refuses to offer any assistance (Kant 1985, 423). The maxim of refusing help to those in need could be coherently thought—but not willed—as a law of nature “since instances can often arise in which he would need the love and sympathy of others, and in which he would have robbed himself, by such a law springing from his own will, of all hope of the aid he desires” (Kant 1985, 423).

Setting aside any concern for the details of Kant’s analyses, we can clearly identify Kant’s four-step algorithm of moral judgment using the Principle of Universalizability as follows:

<b>Step 1</b>	<b>Step 2</b>	<b>Step 3</b>	<b>Step 4</b>
Determine the maxim of the action under consideration.	Universalize the maxim into a law of nature.	Is there a contradiction of <i>thought</i> ? (Perfect duty)	Is there a contradiction of <i>will</i> ? (Imperfect duty)

So after identifying and universalizing the maxim of action, perfect and imperfect duties are determined by a simple, purely logical process of detecting contradictions. An action is a perfect duty if the contrary action “cannot even be *thought* a universal law of nature without contradiction” (Kant 1985, 424). An action is an imperfect duty if the contrary action can be thought as universal law, but not *willed* as such “because such a will would contradict itself” (Kant 1985, 424). Notably, the only difference in the algorithm used in these four examples is that the first two do not bother with the question of conflicts of will in Step Four, as a conflict of thought has already been identified in Step Three.

Perhaps the greatest challenge to the outward simplicity of this algorithm of moral judgment lies in the substantial difficulties of formulating the maxim of action. Philip Pettit aptly summarizes the basic problem thusly:

Kant just assumes that there is a unique maxim, or a uniquely suitable maxim, associated with any option... But that assumption is hard to sustain. There's no known method or algorithm whereby we can be reliably directed to the maxim upon whose universal prescribability the rightness of the option depends. And there's no plausibility in the claim that we can just tell in every case what the relevant maxim is: that innate or culturally shaped intuition can do the job (Pettit 1997, 134).

In particular, the maxim of an action can be formulated with a whole range of details omitted or included—and those details can themselves be formulated in any number of ways. Kant himself provides us with one such problematic example of abstract and concrete maxims in his discussions of false promises. In the first discussion, he uses the ultra-abstract maxim “May I, when in distress, make a promise with the intention not to keep it?” (Kant 1985, 402). But later, the maxim is the far more concrete “When I believe myself to be in need of money, I will borrow money and promise to repay it, although I know I shall never be able to do so” (Kant 1985, 422). In this case, the difference does not impact the duty generated by the universalization algorithm, but not all such examples are so easy. In considering the morality of lies to murderers, for example, Kant is clearly relying upon a very broad maxim along the lines of “I shall lie when it is expedient for me to do so” in order to justify his claim that lying in such cases does “wrong to men in general,” not just to murderers, criminals, and others without the disputed “right to the truth” (Kant SRTL, P6, P3). However, a more concrete formulation of the maxim, perhaps “I shall lie to protect innocent people from death at the hands of murders,” can be coherently universalized, as surely the occasional lie to the rampaging murderer would not undermine trust in utterances generally, although perhaps rampaging murderers would have greater reason for caution. So in this case, the particular formulation of the maxim makes all the difference in the outcome of the algorithm of moral judgment, as one formulation forbids deliberate falsehood in all circumstances, while the other permits it in some.

More generally, Marcia Baron notes that various formulations of the maxim of action can suffer from both false positives (in which “a contradiction appears yet the maxim seems clearly to be permissible”) and false negatives (in which “no contradiction is detected, yet the maxim seems clearly to be unacceptable”) (Baron 1997, 71). The temptation here (which Baron resists) is to simply “restate the maxim... so as to yield the [desired] result” (Baron 1997, 73). However, this solution is deeply anti-Kantian, as it elevates our very fallible moral intuitions over pure reason. We need—but do not have—an independent method of determining the details that ought to be included or excluded in our maxims. Interestingly enough, Marcia Baron does not regard this issue as “a grave problem” because she sees Kantian ethics as one of “self-scrutiny” in which “we need to reflect on just what it is that we are doing, rather than apply a simple, straightforward, fairly mindless test” (Baron 1997, 75). Although this is a plausible and compelling solution, it is not consistent with Kant’s basic aim of constructing a science of ethics.

The process of moral judgment associated with the Principle of Humanity is even simpler than that of the Principle of Universalizability—and without the difficulties of formulating the maxim of action. As before, Kant’s four standard examples reveal the basic algorithm of judgment. In the suicide and false promises examples, the proposed actions contradict “the idea of humanity as an end in itself”—and thereby identify refraining from suicide and truthfulness as perfect duties (Kant 1985, 429). The imperfect duties of developing talents and helping others, in contrast, do not reveal any contradiction with the idea of humanity as an end in itself—but only a lack of harmony with it (Kant 1985, 430). Kant also regards the division between perfect and imperfect duties as a matter of preserving humanity versus furthering it (Kant 1985, 430). Thus the Principle of Humanity generates this simple two-step process of moral judgment:

Step 1	Step 2
Does the action <i>contradict</i> the idea of humanity as an end in itself? ...Or...Is the action inconsistent with the <i>preservation</i> of humanity as an end in itself? (Perfect duty)	Does the action <i>harmonize</i> with the idea of humanity as an end in itself? ...Or... Is the action inconsistent with the <i>furtherance</i> of humanity as an end in itself? (Imperfect duty)

The absence of any reference to the maxim of action in this algorithm is a mixed blessing. Obviously, it thereby avoids the difficulties inherent in formulating such maxims, as discussed above. But at the same time, it renders Kant’s inferences from particular actions to general duties rather mysterious. So in the case of suicide, for example, Kant seems to infer the duty “I cannot dispose of man in my own person so as to mutilate, corrupt, or kill him” directly from the Principle of Humanity (Kant 1985, 429). The murkiness of this inference is only exacerbated by Kant’s qualification that this duty does not forbid amputation due to medical necessity or taking risks in the process of saving one’s own life. On what grounds, we justly wonder, are such actions excluded from the duty?

More interesting are the differences between the duties generated by the Principle of Universalization versus the Principle of Humanity. The duty to refrain from suicide generated by universalization, for example, says nothing about the impermissibility of mutilation or corruption of the body. More significantly, some actions sail through one algorithm while obviously failing the other. For example, the universalization of the maxim “Slaves should obey their masters” generates no contradictions of either thought or will, although a master-slave relationship is clearly inconsistent with the idea of humanity as an end in itself. Such examples show that the various versions of the Categorical Imperative cannot be equivalent.

Looking broadly, perhaps the most remarkable feature of Kant’s algorithmic conception of moral judgment is the absence of any concern for empirical facts. The particular details of a situation, the differences between and preferences of individuals, and even the likely consequences of action are all irrelevant to moral judgment. Moral decisionmaking involves no negotiation between competing concerns and values—and thus no substantial skills of moral judgment cultivated and honed by life experience. Kant explicitly confirms this interpretation as his view of moral judgment in a discussion of false promises and the Principle of Universalization. He writes,

I do not, therefore, need any penetrating acuteness to discern what I have to do in order that my volition may be morally good. Inexperienced in the course of the world, incapable of being prepared for all its contingencies, I only ask myself: Can I will that my maxim become a universal law? If not, it must be rejected, not because of any disadvantage accruing to myself or even to others, but because it cannot enter as a principle into a possible enactment of universal law, and reason extorts from me an immediate respect for such legislation (Kant 1985, 403).

In fact, Kant’s only comment in the *Foundations* on the practical skills of moral decisionmaking is a moderate concession that “a power of judgment sharpened by experience” helps us to “decide in which cases [the a priori moral laws] apply” and to “procure for them access to man’s will and provide impetus for their practice” (Kant 1985, 389). As such, neither experience nor skill plays any role in the actual process of moral judgment for Kant.

Unsurprisingly, modern Kantians often distance Kantian ethics from Kant’s own algorithmic moral judgments. In the “Translators Introduction” to the *Foundations*, for example, Lewis White Beck replies to a number of common criticisms of Kant, including that

Kant minimizes the role of practical intelligence in life, and instead of calling for more facts and acknowledging the more tentative character of our difficult decisions, he requires us only to apply mechanistically a simple logical rule (Beck 1990, xx).

Beck acknowledges the “kernel of truth” in this critique, noting that Kant “often oversimplifies making a moral decision” and citing Kant’s comment about the lack of need for “any penetrating acuteness” (Beck xx). By way of substantive response, Beck claims that Kant’s examples are “simple cases,” but “in real life where duties seem to conflict with one another (and not just with out passions and wishes) things are not so simple” (Beck xx). This reply seems to miss the mark in two ways. First, Kant does not recognize the possibility of actually conflicting duties—nor is he impressed by the obvious conflict of apparent duties in the case of lying to murderers (Kant SRTL). Even in that difficult and horrible situation, the moral course of action is clear and unambiguous to Kant. Second, Kant’s analyses of the four standard examples are simplistic, but the examples themselves are not. They concern exactly the sort of ordinary and everyday moral questions that an ethic must answer. As such, Kant’s selection of examples is not the problem; his quick and easy analysis of them is—and that is precisely the objection, so it cannot also be the response.

Kant’s algorithms of moral judgment, in their rationalistic simplicity, express the same desire for a science of ethics found in his moral standard. But once again, all is not as clean, neat, and simple as Kant would have us believe. The outward minimalism of the four-step and two-step algorithms conceals both deep complexity and substantial inconsistency. Thus neither the moral standard nor the process of moral judgment developed by Kant is capable of transforming ethics into a science.

### **Moral Standard in Mill: Collective Happiness**

In sharp contrast to Kant’s single-minded focus upon motives, the utilitarian ethic advocated by Mill is concerned with the consequences of action, particularly the consequences for the collective happiness of humankind. The first principle and moral standard of Mill’s ethics is the General Happiness Principle, which states that “actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness” (Mill 1987, 278). Happiness, Mill clarifies, is simply “pleasure, and the absence of pain,” while unhappiness is “pain, and the privation of pleasure” (Mill 1987, 278). And pleasures, Mill further insists (although not without difficulty), can be ranked on the basis of quality in addition to quantity. (Mill 1987, 279). Perhaps most significantly, the General Happiness Principle is—as its name implies—wholly impartial, requiring moral agents to weight the happiness of all people equally (Mill 1987, 288). Thus Mill insists that “the standard is not the agent’s own happiness, but the greatest amount of happiness together” (Mill 1987, 282). Or, quoting Bentham, “everybody to count for one, nobody for more than one” (Mill 1987, 336).

Again in contrast to Kant, Mill does not attempt a rigorous proof of the General Happiness Principle. Instead, he argues that “questions of ultimate ends are not amenable to direct proof” because “whatever can be proved to be good, must be so by being shown to be a means to something admitted to be good without proof” (Mill 1987, 275). Mill, however, does not ask us to adopt his proposed final end of collective happiness on “blind impulse” or “arbitrary choice,” instead offering “considerations... capable of determining the intellect either to give or withhold its assent to the doctrine” (Mill 1987, 276). In keeping with Amartya Sen’s helpful breakdown of utilitarianism into the conjunction of consequentialism, welfarism, and sum-ranking, we can divide Mill’s arguments for the General Happiness Principle into the three basic elements of concern for consequences, for happiness, and for impartiality (Sen 1987, 39).

Mill does not directly argue for his narrow focus upon the consequences of action, although his contention that even Kant's ethics cannot completely ignore such consequences, if true, would lend some partial support to this idea. According to Mill's interpretation of the Categorical Imperative, the failure of the universalization of some maxims but not others is not the result of "any contradiction, any logical (not to say physical) impossibility," but rather "that the *consequences* of their universal adoption would be such as no one would choose to occur" (Mill 1987, 275). Mill is certainly right that a world where people offer false promises when expedient, for example, is both metaphysically and physically possible—and perhaps even actual. But this misses Kant's basic point that the effectiveness of false promises logically depends upon the maxim *not* becoming universal law. As such, the contradiction is in the agent, not in external metaphysical or physical circumstances. Even if we grant the accuracy of Mill's analysis of the Categorical Imperative, he has at best shown that ethics must have some concern for consequences, but certainly not necessarily the paramount concern found in utilitarianism. Generally speaking, Mill tends to presume without argument that consequences are all that matter. Or, as he says in a footnote on the differences between motives and intentions, "when [something] makes no difference in the act, [it] makes none in the morality" (Mill 1987, 290).

Mill does, however, offer a more direct argument for the utilitarian concern for happiness. In a generally Aristotelian vein, Mill claims that happiness is the self-evident ultimate end of all people, as it is actually desired as an ultimate end and "all other things [are] only desirable as a means" to it (Mill 1987, 307). Mill tempers this strong "means only" claim in his discussion of virtue, arguing that virtue is not only a means to happiness but also constitutive of it (Mill 1987, 309). Nevertheless, virtue does not compete with happiness for the title of ultimate end, because, as the all-consuming principle in this schema, happiness remains the *ultimate* ultimate end (Mill 1987, 311). Overall, Mill's argument here is neither as detailed nor as persuasive as Aristotle's on this issue. More importantly, Mill's focus on utility as happiness rather than welfare is unfortunate, as its subjectivity ensnares him in a number of significant problems, including his distinction between higher and lower pleasures.

Sloppily injected into the justification for happiness as the ultimate end of individuals is Mill's far more dubious argument for impartiality between the happiness of all people. Mill writes,

No reason can be given why the general happiness is desirable, except that each person, so far as he believes it to be attainable, desires his own happiness. This, however, being a fact, we have not only all the proof of which the case admits of, but all which it is possible to require, that happiness is a good: that each person's happiness is a good to that person, and the general happiness, therefore, a good to the aggregate of all persons (Mill 1987, 307-8).

Consequently, Mill claims, the general happiness is the "criteria of morality" (Mill 1987, 308). This argument is commonly regarded as a flagrant example of the fallacy of composition, as we cannot logically infer that a "general happiness" exists as "a good to the aggregate of all people" on the grounds that each person's own happiness is a good to him as an individual. To claim that would be like inferring that a "general wife" exists as "a good to the aggregate of all happily married husbands" because individual wives are goods to their corresponding individual husbands. However, Mill disputes this interpretation in a private letter, writing that

As to the sentence ... [where] I said that the general happiness is a good to the aggregate of all persons I did not mean that every human being's happiness is a good to every other human being, though I think that in a good state of society and education it would be so. I merely meant... to argue that since A's happiness is a good, B's a good, C's a good, etc., the sum of all these goods must be a good (Mill 1987, 1972, 1414).

Unfortunately Mill's clarification only lessens the severity of the fallacy of composition without eliminating it entirely, as we still must wonder for whom the collective good is a good at all: The

individuals? The collective? In fact, neither will do. Individuals have only been offered a compelling reason to recognize that other people value their own individual happiness, not an independent reason to value or promote that happiness themselves. And the collective is not a moral agent capable of valuing or of imposing moral obligations upon its constituent parts (at least not in any simple way).

Mill does attempt to justify impartiality through a slightly more plausible Humean appeal to the “powerful natural sentiment” of “desire to be in unity with our fellow creatures” (Mill 1987, 303). As social creatures, Mill argues, we must recognize that “society between equals can only exist on the understanding that the interests of all are to be regarded equally”—and that any other foundation gives rise to a society of “master and slave” (Mill 1987, 304). Unfortunately for Mill, the appeal of this argument—and the tale of the moral development that follows—relies upon his covert shift to an impartialist perspective overlooking the whole of society so as to pump our intuitions about the morality of impartialism from the very different perspective of an individual within that society. But accepting the proposition that all people ought to be equal before the law does not require us to grant that all people ought to be equal in our moral deliberations. In addition, Mill’s association of partialism with slavery is bizarre, as a woman who, acting on her partialist feelings, spends \$20 on flowers for her mother rather than donating the money to a charity for the poor is not in any way promoting, replicating, or acting out the injustice of slavery. She is not denying the basic moral worth of people outside her sphere of concern, only choosing not to actively promote their interests. So here too, Mill gives us little reason to adopt the impartialist moral standard of the General Happiness Principle.

In the end, Mill’s justification of the General Happiness Principle is largely an appeal to intuition. He claims that “whatever steadiness or consistency” moral doctrines have achieved can be attributed to the “tacit influence” of the Greatest Happiness Principle. Even “the moral doctrines of those who most scornfully reject its authority” have been shaped by it. Furthermore, Mill’s arguments about justice and virtue seem aimed at reconciling our common-sense moral intuitions with utilitarianism. Kant, in contrast, is perfectly happy to bite the bullet when our ordinary moral intuitions clash with the demands of duty. Mill’s General Happiness Principle is thus accordingly subject to critique on the basis of moral intuitions in way that Kant’s Categorical Imperative is not. And indeed, the principles and prescriptions of utilitarianism do often conflict with our basic moral intuitions. For example, the doctrine of double effect indicates that differences in intentions are relevant to the moral evaluation of two acts with identical consequences, which utilitarianism must deny. Utilitarianism can also rather easily justify morally repugnant institutions like slavery and other violations of rights if the general happiness would be benefited thereby. Even worse, the impartialism of utilitarianism not only requires us to include the happiness of murderers, rapists, and dictators in our moral calculus, but also regard their happiness as “equally desirable” as the happiness of their victims and other moral innocents (Mill 1987, 336ft).

Generally speaking, the moral standard of the General Happiness Principle requires a far more sturdy foundation than Mill’s unconvincing arguments and weak appeals to intuition. Mill’s attempt at an empirical science of ethics, at least insofar as the moral standard is concerned, is simply inadequately justified.

### **Moral Judgment in Mill: The Mathematics of Utility Comparison**

With this picture of Mill's utilitarian moral standard in mind, we can now turn our attention to his method of using that standard in everyday moral decisions. The basic process of utilitarian moral judgment is fairly simple and clear—at least in theory. We must:

1. Evaluate the impact of possible actions upon the happiness of all affected people.
2. Sum up the utilities of all people for each of these possible actions.
3. Choose the action that maximizes collective utility.

Beneath the simplicity of this algorithm lies a whole host of significant practical difficulties.

First and foremost, the fact that happiness is a private and subjective state creates substantial difficulties measuring and comparing the happiness of other people (particularly distant people) with any degree of accuracy. More publicly accessible factors, such as common sources of happiness (e.g. wealth and friendship) and expressions of happiness (e.g. exclamations and body language), are strongly influenced by individual preferences and temperaments. The particular danger of ignorance is that we are liable to paternalistically impose our preferences upon others, thereby potentially undermining individual uniqueness and cultural differences. Additionally, our actions may have substantial secondary, tertiary, and other remote effects that, if included in our calculations, make them exponentially more complex. But failing to include them vitiates moral responsibility for the indirect effects of our actions, even when foreseeable and/or foreseen. For example, a parent who negligently leaves a loaded firearm on the counter might say “Well, I just left the gun on top of the counter. It wasn't harming anyone there. Sure, my kid picked it up and accidentally shot himself. But that's a failure of his utilitarian calculus, not mine. I don't complicate my calculus with remote consequences.” If we are to include remote consequences in the utilitarian calculus, however, we must have some method of determining which to include and which to omit.

Mill attempts to avoid the burdens of utilitarian calculation through the use of “intermediate generalizations” to guide day-to-day action. After all, argues Mill, utilitarianism does not commit us to the impossible task of “calculating and weighing the effects of any line of conduct on the general happiness” for every action at the moment of choice (Mill 1987, 295). Instead, as with most other moral theories, we have sufficient time in advance to “learn by experience the tendencies of action” and thus draw “general conclusions from the experience of human life” (Mill 1987, 295, 297). Or, put more forcefully:

People talk as if the commencement of this course of experience had hitherto been put off, and as if, at the moment which some man feels tempted to meddle with the property or life of another, he had to being considering for the first time whether murder and theft are injurious to human happiness (Mill 1987, 295).

As such, we are like the sailor who goes out to sea with navigational calculations already performed and recorded in his nautical almanac, for as “rational creatures,” we “go out upon the sea of life with [our] minds made up on the common questions of right and wrong, as well as on many of the far more difficult questions of wise and foolish” (Mill 1987, 296-7).

In these advance deliberations, Mill recommends provisionally accepting the basic principles of “the received code of ethics” and then gradually updating that code with any new knowledge of “the effects of actions on the general happiness” (Mill 1987, 296). In all cases, our intermediate generalization must serve the General Happiness Principle (Mill 1987, 297). Mill also cautions that these general rules “cannot be so framed as to require no exceptions” such that some actions are be “always obligatory” and others are “always condemnable” (Mill 1987, 297). Our intermediate generalizations will include “a certain latitude, under the moral responsibility of the agent, for accommodation to peculiarities of circumstances” (Mill 1987, 297). However, given the amazing variety particulars that might impact the general happiness, we have some

reason to worry that our intermediate generalizations will become so weighted down with exceptions as to be unable to guide choices and actions efficiently.

The most glaring inconsistency in Mill's conception of moral judgment is his argument that we need only be concerned with the promotion of local happiness. He claims that we need not "fix [our] minds upon so wide a generality as the world, or society at large" but rather merely upon "particular persons" (Mill 1987, 290-1). Very few people, Mill argues, are capable of assuming the role of the "public benefactor" (Mill 1987, 291). Those who are should "consider public utility" but the rest of us need only concern ourselves with "the interest or happiness of some few persons" (Mill 1987, 291). However, particularly in light of diminishing marginal utility, the General Happiness Principle does seem to require non-local action, particularly for those of us living in the comforts of relative prosperity. After all, my time, energy, and money can only slightly benefit my friends and acquaintances, as most of them also already live comfortably enough. If instead directed toward helping the desperately poor, sick, and otherwise needy, my resources would have a substantially greater impact upon the general happiness.

Along these lines, Michael Slote notes that utilitarianism seems to generate

the moral duty to abandon one's professional or business career, if (one thinks) one can do more good in the world by doing to help the sick and starving at home and abroad than one could by pursuing one's career, being kind of those near and dear to one, and even given generously to charity (all while living in relative comfort (Slote 1997, 190).

The strict adherence to the General Happiness Principle thus requires "more self-sacrifice on the part of (many) moral agents than, intuitively, it seems fair to require of them" (Slote 1997, 190).

Mill's vision of moral judgment does seem to avoid the worst of the utilitarian calculation problems by appealing to intermediate generalizations. But in doing so, the simple, scientific-seeming algorithm of calculating and adding utilities is obscured in complexity.

### **Ethics as Science or Art**

With our critical examination of both Kant's and Mill's ethics behind us, we can now clearly see that they have not lived up to the hope of ethics as a science. Neither philosopher's moral standard is well-grounded, and neither philosopher's process of moral judgment is as clear and simple in practice as in theory. In short, the necessary, universal, and simple truths that characterize the sciences are not found in either Kant or Mill.

In contrast to both these philosophers, Aristotle is perfectly satisfied with and even requires a fair amount of art in his ethics. In the *Nicomachean Ethics*, he warns us "not look for precision in all things alike, but in each class of things such precision as accords with the subject-matter, and so much as is appropriate to the inquiry" (Aristotle, 1098a). As an inherently practical rather than theoretical subject, we should approach ethics like the carpenter who measures right angles well enough to suit his practical purposes rather than like the geometer who is concerned with the abstract and universal nature of right angles (Aristotle 1098a). To do otherwise would subordinate the "main task... to minor questions" (Aristotle 1098a).

In Aristotle's substantive ethics, much of the imprecision emerges from the recognition of and respect for individual differences between people, knowledge of the complexity of the particulars in any moral situation, and the use of largely subconscious skills, habits, and feelings to guide moral judgment. Such features are strengths, not weaknesses. And this complexity, even given an inability to consciously articulate and teach all the relevant principles and method, does not make Aristotle's ethics particularly difficult to either understand or practice. On the contrary, Aristotle's basic project of making virtue second nature through cultivated habits and

dispositions means that living a moral life will likely be easier than in Kantianism or utilitarianism.

Of course, we may not wish to adopt the substance of much of Aristotle's moral theory. Nevertheless, Aristotle reminds us that a robust, livable ethic need not rely upon the model of science, of simple universal rules and algorithms, but instead may include all the complexity of an art.

## References

- Aristotle. *Nicomachean Ethics*. Translated by W.D. Ross, revised by J.O. Urmson.
- Baron, Marcia. 1997. "Kantian Ethics" in *Three Methods of Ethics*. Oxford: Blackwell Publishers.
- Beck, Lewis White. 1985. "Introduction" in *Foundations of the Metaphysics of Morals*. 2<sup>nd</sup> ed. New York: MacMillan.
- Kant, Immanuel. 1985. *Foundations of the Metaphysics of Morals*. Translated by Lewis White Beck. 2<sup>nd</sup> ed. New York: MacMillan.
- Kant, Immanuel. 2003. "On A Supposed Right To Tell Lies From Benevolent Motives." Translated by T.K. Abbot. Available at <<http://core.ecu.edu/phil/mccartyr/kant/entry3.htm>>.
- Mill, J. S. 1972. *Later Letters*. Volume XVI. Toronto: University of Toronto Press.
- Mill, J. S. 1987. "Utilitarianism" in *Utilitarianism and Other Essays*. Edited by Alan Ryan. London: Penguin Books.
- Paton, H. J. 1965. *The Categorical Imperative*. London: Hutchinson & Co.
- Pettit, Philip. 1997. "The Consequentialist Perspective" in *Three Methods of Ethics*. Oxford: Blackwell Publishers.
- Ross, David. 1954. *Kant's Ethical Theory*. London: Oxford UP.
- Ryan, Alan. 1987. "Introduction" in *Utilitarianism and Other Essays*. Edited by Alan Ryan. London: Penguin Books.
- Sen, Amartya. 1987. *On Ethics and Economics*. Oxford: Blackwell Publishers.
- Slote, Michael. 1997. "Virtue Ethics" in *Three Methods of Ethics*. Oxford: Blackwell Publishers.
- Wilkinson, Will. 2001. "Kant on Happiness In Part I of Groundwork for the Metaphysics of Morals." Available at <<http://willwilkinson.net/essays/kant.html>>.