Phil 4: Intro Ethics April 26, 2021
Handout #2 Prof. Aaron Z. Zimmerman

Mill's Utilitarianism

I. Three Branches of Ethics

1. *Normative Ethics*: the study and development of systems of right and wrong—systems of rules, principles or procedures for figuring out what one should do and should not do (morally speaking).

> Examples of normative ethical systems: Utilitarianism, Kantian or Deontological Ethics, Virtue Ethics, Religious Moralities (e.g. Judeo-Christian law, Islamic code, etc.), Transcendentalism

2. *Ethical Motivation* (or Moral Psychology): the study of the source of moral obligation—an evaluation of answers to the questions "Why do people act morally (when they do)?" and "Why should we be Moral?"

> Examples of theories of moral motivation: a) Religious—(a) normative: we should be moral because God wants us to be moral, or loves what is right; (b) psychological: people sometimes or always act morally (even when they know they are sufficiently clever or powerful to avoid earthly punishments) out of fear of God's wrath or a desire for God's love; b) Teleological— (a) normative: we should be moral because the function of people is to be moral, people are "malfunctioning" when they act immorally; (b) psychological: people will act morally so long as their cognitive and/or affective faculties are not (in some biologically defined sense) impaired; c) Rational—(a) normative: we should be moral because (Hobbes) in the long run immorality is contrary to our own "selfish" interests or (Kant) immoral motives involve some sort of inconsistency or incoherence; (b) psychological: we act morally when we (Hobbes) deliberate clearly giving proper weight to our long-term self-interest, or (Kant) choose rules for action in a coherent or consistent manner

3. *Meta-ethics*: the study of *moral epistemology* and *moral metaphysics*. (a) Moral epistemology is the study of moral knowledge and the justification of our moral beliefs—an evaluation of answers to the questions: Do we have any moral knowledge? Are any of our moral beliefs rationally held? Can we rationally resolve disagreements on moral matters? Can we rationally resolve conflicts internal to our common sense moral views? How can we rationally extend our moral views to matters on which we currently find ourselves with no settled opinion? How do we know right from wrong? (b) Moral metaphysics is the study of the nature of moral phenomena — an evaluation of answers to the questions: Are there facts about what is wrong and what is right? If there are such facts, what *makes* something wrong or right? How did certain things *come to be* good and other things *come to be* bad? What is the relation between *moral facts*—e.g. facts about what is good or bad—and *social scientific facts* about the physical world—e.g. biological, psychological, sociological and economic facts about (a) the kinds of actions, laws and institutions that tend to promote happiness, pleasure and satisfaction or misery, pain and dissatisfaction (however measured) (b) facts about those actions, laws and institutions that tend not to do so, (c) facts about the nature of suffering and happiness themselves, and (d) facts about what people actually think, feel and do? Do the social scientific facts entail the moral facts? Or are value-laden facts wholly "distinct" from value-neutral facts? If value-laden claims can be distinguished from value-neutral claims, under what conditions can we reasonably or cogently infer an "ought" from an "is"? Are there any moral facts that are universal in scope, or are all acts right at some places and times and wrong at others?

> Examples of meta-ethical positions: Expressivism, Nihilism, Projectivism, Constructivism and various other forms of Anti-Realism, Realism (Natural and Non-natural), and Relativism.

II. Utilitarianism

A. Utilitarianism is commonly interpreted as a normative ethical view: it consists in the Principle of Utility, which is a *First Principle* of morality. The Principle of Utility is usually formulated as a claim about what we should do, but Mill advances it as a claim about which actions are right and which wrong.

> **The Principle of Utility** ("the greatest happiness principle"): actions are right in proportion as they tend to promote happiness; wrong as they tend to produce the reverse of happiness.

Mill's initial definition of 'happiness': "by 'happiness' is intended pleasure and the absence of pain; by 'unhappiness', pain and the privation of pleasure."

B. In *Utilitarianism*, Mill addresses questions from all three branches of ethics described above. Here's an overview:

* In **chapter 1** he explains what a *first principle* of morality is and argues that morality must have a first principle.

* In **chapter 2** he argues that the principle of utility is the first principle of morality. He then does two more things: first he tries to explain the principle of utility by giving a relatively sophisticated account of happiness; then he defends the claim that the principle of utility is the first principle of morality from some objections. Thus, in chapter 2, Mill is concerned with advancing and defending a position in normative ethics.

* In **chapter 3** Mill turns to questions of moral motivation, and asks, "What are the motives or reasons we have for obeying the principle of utility?" That is, why *do* we act benevolently (when we do) and why *should* we act so as to maximize the happiness produced by our actions? What arguments can be given to convince someone to live by utilitarian norms or rules when she does not already live in this way or see the wisdom of living in this way?

* In **chapter 4** Mill addresses meta-ethical concerns. He tries to prove that the principle of utility is *true*—that it is a *fact* that we should act so as to promote happiness, and he tries to say something about what sort of fact this is and how it is connected to less value-laden sorts of facts—here facts about what normal people actually desire when they are thinking clearly.

* In **chapter 5** Mill returns to normative ethics and addresses the most important kind of objection to Utilitarianism—the claim that it cannot account for considerations of justice or rights. This criticism is often sharpened by arguing that Utilitarianism implies that we are sometimes permitted or even obligated to violate the rights of innocent people. Mill tries to argue (a) that some of intuitions about justice and rights are not to be trusted, but (b) adopting the utilitarian ethic is compatible with retaining certain beliefs about justice (i.e. those not grounded in a thirst for revenge and other retributivist emotions). He also gives us a way to distinguish the utilitarian ideal of producing happiness from the utilitarian conception of obligation which is defined in terms of the utility of imposing sanctions or punishments of various sorts.

III. Chapter 1: What is a first principle? Why must morality have one?

A. Phrases Mill uses interchangeably with 'first principle': 'the criterion of right and wrong', 'the summum bonum', 'the foundation of morality', a 'test of right and wrong', and 'an ultimate standard'.

Mill's argument for the necessity of a first principle of morality:

B. Most sciences—or branches of inquiry—are not based on first principles. Contrary to some appearances algebra is not based on axioms: "algebra derives none of its certainty from what are commonly

taught to learners as its elements, since these. . .are as full of fictions as English law, and of mysteries as theology."

C. But morality and its development is a "practical art." A morality doesn't just tell us how things are, and how they will be, it tells us how we should make things be. It provides us with rules of action, and, according to Mill, *rules of action tell us how to act so as to achieve some end or goal*. Though the moral rules or precepts we live by are distinct from the laws that are operative in our state or nation, Mill is arguing that *morality and law are structurally similar*. They are the means our ancestors have crafted to achieve various goals. But, Mill asks, how can we evaluate competing rules of action unless we know what goal (or goals) we are trying to achieve?

"When we engage in a pursuit, a clear and precise conception of what we are pursuing would seem the first thing we need, instead of the last we are to look forward to. A test of right and wrong must be the means, one would think, of ascertaining what is right or wrong, and not a consequence of having already ascertained it."

Mill then admits that we needn't have a first principle if we have some way of **ranking** principles in order of their importance or force. He then makes a curious remark about how we would know we had uncovered such a first-principle. "There ought either to be some fundamental principle or law at the root of all morality, or, if there be several, there should be a determinate order of precedence among them, and the one principle, or the rule for deciding between the various principles when they conflict, ought to be *self-evident*."

[Notice here the slide from a) a *decision procedure*—like ranking principles, and always following the principle with a higher rank, to b) a *moral principle* by which one decides between the various competing rules.]

Question: Is the principle of utility a good candidate for a self-evident truth? Is it a principle that you can tell is true simply by understanding its meaning? What is the connection between "goodness" and "happiness"? Is happiness the only thing that is intrinsically good? Is happiness a good candidate for the "goal" in light of which we should craft our moral norms and laws?

IV. Chapter 2: Describing and Defending the Principle of Utility

The principle of utility says that acts are right insofar as they promote happiness and prevent suffering and wrong when they promote pain and diminish happiness. But we need to know at least three things if we are to have a more determinate grasp of the principle than this statement provides.

Question 1. "Happiness" is a somewhat vague term. So what is happiness?

Question 2. Once we've further defined "happiness" we can ask: Whose happiness matters? Whose happiness should we consider when we're applying the principle of utility? How much weight should we give to the happiness of the various people affected?

Question 3. What is meant by 'promotion'? Are we simply required to take steps to increase happiness and diminish suffering or must we seek to maximize happiness and eliminate suffering? Should we seek to maximize aggregate happiness or average happiness?

In regard to #2, Mill says that the happiness of (a) "all sentient creatures," should be counted for something, and says that the pleasures and pains of (at least) each person should be counted equally.

> ***"The happiness which forms the utilitarian standard of what is right conduct is not the agent's own happiness but that of all concerned. As between his own happiness and that of others, utilitarianism requires him to be as strictly impartial as a disinterested and benevolent spectator."***

This idea of "equal weight" gives rise to the most challenging aspects of utilitarianism. Notice here that if one's own happiness is not to be valued above the happiness of others, it can hardly be argued that the happiness of one's loved ones can be counted as more important than the happiness of others. Some utilitarians argue, on this basis, that we shouldn't value our children over the children of strangers to the extreme extent that most of us do. See the true stories of Zell Kravinsky who donated a kidney to a stranger, and the "effective altruists" described by Larissa MacFarquhar in her book Strangers Drowning. Just to give you a sense: one couple adopted 20 foster kids even though they had 2 biological children and new their biological kids would get less attention (and likely be less happy?) as a result.

The idea expressed by Mill's quote above suggests that we should not favor anyone over anyone else in our efforts to maximize happiness and that we should always assume this policy of neutrality. Some people draw this conclusion from Peter Singer's landmark essay "Famine, Affluence and Morality," which is posted on the course website as recommended reading. In that essay, Singer argues that benefiting yourself in a minor way by spending $150 on a new pair of shoes is immoral when you know you could use that same money to save the life of someone suffering from famine or disease, and he generalizes from this to argue that we should give to others until doing so would make us worse off than them. (He has since defended more modest proposals.) Is neutrality of this sort Singer recommends always appropriate? Is it true that "as between [your] own happiness and that of others, utilitarianism requires [you] to be as strictly impartial as a disinterested and benevolent spectator."

<u>How does this chapter 1 claim jibe with what Mill says in chapter five about neutrality</u>?

"Impartiality, however, does not seem to be regarded as a duty in itself, but rather as instrumental to some other duty; for it is admitted that favour and preference are not always censurable, and indeed the cases in which they are condemned are rather the exception than the rule. A person would be more likely to be blamed than applauded for giving his family and friends no superiority in good offices over strangers, when he could do so without violating any other duty; and no one thinks it unjust to seek one person to another as a friend, connection or companion."

Further questions: Admittedly, in this passage Mill is describing our intuitive views about justice rather than articulating or defending his utilitarian ethic, which is intended as a modification of the customary morals accepted by most (if not all) of the audience to whom Mill's book was addressed. But Mill never says this intuitive limitation on neutrality is mistaken in any way. And the question remains: Is complete neutrality with regard to the wellbeing of those your actions affect, compatible with your enjoying and sustaining friendships? It seems that Mill has a different idea of the place for neutrality. Mill's father James and Jeremy Bentham argued that legislators ought to be neutral as between the happiness of their constituents when formulating and voting on laws. Mightn't Mill embrace a similar view with regard to moral rules or social norms: *when deciding on these general prohibitions and allowances* (in our more reflective moments) we ought to have as our goal the maximization of happiness and remain neutral as to "where" this happiness is likely to be experienced.

In regard to #3, Mill says the ultimate end is the *maximization* of *aggregate* happiness.

In regard to #1 Mill develops a qualitative view of pleasure and pain according to which *some pleasures and pains count more than others*. To compare the relative weight (or importance) of a pleasure Mill suggests the following test,

> "Of two pleasures, if there be one to which all or almost all who have experience of both give a decided preference, irrespective of any feeling of moral obligation to prefer it, that is the more desirable pleasure. If one of the two is, by those who are competently acquainted with both, placed so far above the other that they prefer it, even though knowing it to be attended with a greater amount of discontent, and would not resign it for any quantity of other pleasure which their nature is capable of, we are justified in ascribing to the preferred enjoyment a superiority in quality so far outweighing quantity as to render it, in comparison, of small account."

Mill is confident that the pleasures that this test will deem higher in value (and thus “worth more” in a utilitarian calculus) will be the type-2 pleasures: the pleasures we derive from the use of our “higher” faculties, where higher faculties are those that distinguish people from other animals.

> **Type 1 pleasures**: pleasures we derive from exercising those psychological capacities we share with other animals—e.g. pleasures from eating, drinking, sexual reproduction, play, exploration, etc.
>
> **Type 2 pleasures**: the pleasures we derive from exercising uniquely human capacities—e.g. sentential language, conversation, mathematics, science, business, etc.

“It is an unquestionable fact that those who are equally acquainted with, and equally capable of appreciating and enjoying, both, do give a most marked preference to the manner of existence which employs their higher faculties. Few human creatures would consent to be changed into any of the lower animals, for a promise of the fullest allowance of a beast's pleasures; no intelligent human being would consent to be a fool, no instructed person would be an ignoramus, no person of feeling and conscience would be selfish and base, even though they should be persuaded that the fool, the dunce, or the rascal is better satisfied with his lot than they are with theirs. They would not resign what they possess more than he for the most complete satisfaction of all the desires which they have in common with him.”

> **Three questions**: Is Mill’s prediction about the outcome of such a test correct? Do those who have experienced both prefer type 2 to type 1 pleasures? Is such a test a good way to determine the degrees of pleasure or happiness that should figure in a utilitarian’s application of the principle of utility? Does the qualitative view of happiness undercut the standing of the principle of utility as a *first* principle of morality?

Again, according to Mill, the principle of utility is the first principle of morality. Let’s then call more determinate or less general rules such as “Don’t lie,” “Don’t kill,” and “Don’t steal” **secondary moral rules**. With this distinction in mind, consider the following variations on the principle of utility.

Direct Utilitarianism vs. Indirect Utilitarianism

> **Direct Utilitarianism:** At any given time you should: (a) identify the set of actions open to you at that time, (b) determine which of these actions will produce at least as much happiness as any other action available to you at the time, and (c) perform that action (or one of those actions) that you judge will produce at least as much happiness as would any other course of action available to you at the time.
>
> **Indirect Utilitarianism**: (1) Your actions should conform to a set of secondary moral rules. (2) In quiet moments (or moments of reflection) you should take the time to select (and revise) the set of secondary rules that guide your actions. (3) When selecting or revising the secondary rules from which you will act, you should identify those rules the adoption of which (by you or your community) would lead to the greatest aggregate happiness. (4) You should only adopt a set of rules if you judge that the adoption of those rules (by you or your community) would lead to at least as much happiness as would the adoption of any alternative set of rules.

Mill seems to accept some version of Indirect Utilitarianism in at least this respect: he doesn’t think it’s practical for us to regularly apply the principle of utility when we’re faced with particular decisions in the ordinary course of life. You *need not*—indeed, you *should not*—decide what to do by figuring out which of the actions available to you is likely to maximize aggregate happiness. So according to Mill, there’s a sense in which it is not the case that we should apply the principle of utility to particular cases. We should instead apply it to rules.

> To consider the rules of morality as improvable is one thing: to pass over the intermediate generalization entirely and endeavor to test each individual action directly by the first

> principle is another. It is a strange notion that the acknowledgment of a first principle is inconsistent with the admission of secondary ones. (Mill, 24)

Note that here Mill provides another role for a first principle of morality. A first principle is supposed to allow for *moral progress* or the development of our moral system. It is supposed to provide an antidote to *conservatism*. (Question: what are the other two functions?)

Utilitarian moral theorists are often divided into those that advocate direct utilitarianism and those who advocate indirect utilitarianism. But they are also commonly divided into "act utilitarians" and "rule utilitarians". Though it's pretty clear that Mill advocates indirect applications of the principle of utility rather than its direct use, he might still be an act utilitarian in thinking that what *makes a consequence or state of affairs morally good or worthy of choice* is whether it in fact maximizes utility.

Act Utilitarianism vs. Rule Utilitarianism

> **Act Utilitarianism:** The morally best world (or state of affairs) to create at a given time t is that world (or state of affairs) that contains at least as much happiness any other state of affairs that can be created at t.
>
> **Rule Utilitarianism**: (1) The morally best world (or state of affairs) to create at a given time t is that world (or state of affairs) in which your actions (or the actions of those in your community) conform to the morally best set of secondary moral rules. (2) The morally best set of secondary moral rules are those rules the adoption of which by you (or your community) would lead to at least as much aggregate happiness as would the adoption of any alternative set of secondary rules.

Question: What is it for you or your community to "adopt" a rule?

Answer: It does not involve full compliance, as you can act contrary to rules accepted by most of the people with whom you live. Indeed, on most accounts, you can violate one of your own moral rules, so long as you feel bad about it. But most theorists insist that your adoption of a rule involves a fairly established tendency to conform to that rule and a disposition to feel bad (guilty, remorseful or at least troubled) when you fail to conform to it. Further complexity arises from hypocrisy: you may accept a rule insofar as you judge or shun people who break that rule, even if you don't follow that rule yourself. To understand the relevant phenomenon further reflect on the rules you claim to accept: e.g. "don't kiss on the first date," "look both ways before crossing the street" etc.

Suppose that Mill accepts the view of goodness (or choice-worthiness) we've defined as act utilitarianism above, but that he also advocates an indirect rather than a direct application of the PU, and so embraces the claim we've titled "indirect utilitarianism". In other words, suppose Mill thinks that the best world to create is that world that contains the most aggregate happiness possible, but he thinks we should do our best to follow those of our secondary moral rules that withstand utilitarian scrutiny: we should not apply the principle of utility more regularly than this (much less every time we act or deliberate). Then, ***unless Mill thinks we shouldn't do what would make the world the best it can be*, Mill is committed to the hypothesis that *we will in fact produce more happiness if we act on secondary moral rules than we will if we consciously try to maximize happiness when acting or deciding what to do throughout our days***.

> "According to the Greatest Happiness Principle. . .the ultimate end, with reference to and for the sake of which all other things are desirable (whether we are considering our own good or the good of other people), is an existence exempt as far as possible from pain, and as rich as possible in enjoyments, both in point of quantity and quality. . .This being, according to the utilitarian opinion, the end of human action, is necessarily also the standard of morality; which may accordingly be defined, the rules and precepts for human conduct, by the observance of which an existence such as has been described might be, to the greatest extent possible, secured to all mankind; and not to them only, but, so far as the nature of things admits, to the whole sentient creation."

Questions: How plausible is Mill's mixed view: i.e. **indirect act utlitarianism**? Is direct utilitarianism psychologically possible? Could you adopt it as your ethical guide? What about indirect utilitarianism? Which of our laws and/or norms would we abandon were we to evaluate them in a utilitarian manner?

V. Chapter 3: Moral Motivation

1. Ordinary moral rules bind us because we have been "trained" (by education and opinion) to follow them. But since we were not brought up as utilitarians we do not regard the principle of utility with a "feeling" of obligation. This might lead us to think that we need a reason if we are to accept that we should promote general happiness, but that we do not need a reason "not to rob or murder, betray or deceive" (p. 26). But this is wrong. If we need a reason to follow the one moral principle we need a reason to follow any moral principle.

2. The reasons for following the dictates of ordinary morality are also reasons for adopting the principle of utility. They are both *internal* and *external*:

> *external*: the hope of favor and the fear of displeasure of others (including, possibly, some God).
>
> *internal*: a feeling of conscience "which in properly cultivated moral natures" arises and is strong enough to make immoral action too "painful" to perform.

Questions: Are these good *reasons* for acting morally, or are they only *causes* of moral behavior? Is a sense of obligation (to take care of one's children or stay true to one's boyfriend or girlfriend or respect the property of another person) properly analyzed as aversion to the pain one knows or believes one would experience were one to shirk that obligation? Are there moral motives that cannot be analyzed in terms of a desire for pleasure or an aversion to pain?

3. Mill's reply to the moral skeptic: "Undoubtedly this sanction has no binding efficacy on those who do not possess the feelings it appeals to; but neither will these persons be more obedient to any other moral principle than the utilitarian one." "How can the will to be virtuous, where it does not exist in sufficient force, be implanted or awakened? Only by making the person desire virtue—by making him think of it in a pleasurable light, or of its absence in a painful one. It is by associating the doing right with pleasure, or the wrong with pain, or by eliciting and impressing and bringing home to the person's experience the pleasure naturally involved in the one or the pain in the other, that it is possible to call forth that will to be virtuous which, when confirmed, acts without any thought of either pleasure or pain."

Questions: What is Mill's view of moral training? How does it differ from training pets and other animals through rewards and punishments? Is moral education limited to "conditioning" of this sort? What role does religious instruction play? What role do tales of saints and sinners (and other "morality tales") play in entraining the association of pleasure with brining pleasure to others and pain with brining them pain? What role do reasoning and argumentation play in moral education? How do different cultures and societies answer these questions?

Which kinds of proposition can be treated as premises in an argument for the immorality of an act or the value of a law, rule or plan? Do people disagree about the premises of such arguments? Suppose that a skeptic grants that a given plan or prospective course of action will produce much more misery or suffering than happiness or pleasure, but that this skeptic refuses to conclude that the act is immoral or bad. Does Mill think that such a skeptic is making a mistake? What kind of mistake?

VI. Chapter 4 (Meta-ethics): The proof of Utilitarianism

As utilitarianism is a theory about what is most desirable, or what is desirable "in itself," it does not admit of a proof in the standard sense. Still, "The only proof capable of being given that an object is visible is that people actually see it. The only proof that a sound is audible is that people actually hear it; and so of

the other sources of our experience. In like manner, I apprehend, the sole evidence it is possible to produce that anything is desirable is that people do actually desire it."

Initial Questions: Might someone desire something that is not desirable? Consider a case in which food looks good but tastes bad. (Or someone looks attractive but is actually horrible to spend time with.) You might say something like, "I thought I wanted that, but I was wrong" or (more or less equivalently) "I wanted that because it looked as though it would be enjoyable, but I don't want it now because I now know that it's horrible." Surely, Mill must take this phenomenon into account. Still, he might ask whether you can desire something about which you are **fully and accurately informed** without its being desirable? And if you do equate what is desirable with what you would want upon full information, you might take this to support Mill's claim that your desires are **the best evidence you could have** with regard to what is "really" desirable.

(1) Mill's desire-based characterization of intrinsic goodness: X is intrinsically good or desirable for us just in case we would desire X for its own sake upon full and accurate information about X.

With this claim in mind, consider the conclusion of Mill's proof of utilitarianism, now modified to allow for ignorance-based desires for things that are not really desirable:

(2) **The Conclusion of Mill's Proof:** Happiness is the only thing we desire for its own sake upon full and accurate information**.**

And recall, too, Mill's claim that "happiness" can be defined in terms of pleasure.

(3) **Mill's Hedonistic Conception of Happiness**: Our happiness consists in the type-1 sensory or animal pleasures we get from food, sex and play and the type-2 pleasures we get from exercising our uniquely human capacities for conversation, math, and science, with a predominance of the type-2 pleasures over the type-1.

Remember that Mill characterizes the "happy life" marked by such pleasures as, "not a life of rapture; but moments of such, in an existence made up of few and transitory pains, many and various pleasures, with a decided predominance of the active over the passive, and having as the foundation of the whole, not to expect more from life than it is capable of bestowing."

Now note that Mill's claims (2) and (3) entail:

(4) **A Hedonistic Conception of Final Ends**: The only thing we desire for its own sake upon full and accurate information is pleasure (as described above).

The problem raised by the miser and the monk is that they don't seem to desire pleasure for its own sake, especially when pleasure is characterized in the way Mill has described it. The miser seems to desire money quite apart from the pleasures he might derive from exercising his capacities, and the monk seems to desire virtue quite apart from these same forms of pleasure.

Of course, a miser might take pleasure in just knowing that he has a lot of money. His mere knowledge of his wealth might ease his anxiety, or the thought of his purchasing power might keep a smile on his face. But this is not the kind of miser that raises a problem for Mill's proof. The problem is that some misers seem to value money quite apart from these identifiable pleasures, as they are not made visibly happy by their wealth. (Because they aren't spending their money on fun objects and activities, they aren't deriving the type-1 and type-2 pleasures from it that Mill initially identifies with happiness.) And if a miser fails to display the behavior and affect of someone we would describe as happy (if he is mopey and seems miserable despite his wealth) we would be hard pressed to argue that his money is nevertheless making him happy. Something similar might be said of a stern or visibly miserable priest or monk. Why are they pursuing money and/or virtue when it is obviously not making them happy?

Conclusion: Unless the misers and monks who meet this description are self-deceived about their lives of wealth and virtue (and the relative scarcity of type-1 and type-2 pleasures they're experiencing when leading these lives) Mill has to avoid drawing claim (4) as a conclusion. (He must acknowledge that we have final ends other than pleasure.) And to avoid (4), Mill must either abandon his claim (2) or his claim (3) described above. Mill must either say: (a) that the miser wants something other than happiness for its own sake: possessing money. Or Mill must say: (b) that happiness cannot be equated with pleasure: that the miser's happiness consists in his *merely having money* even when it is fairly obvious that this is not bringing him pleasure.

Perhaps (a) is the more plausible conclusion to draw. After all, we wouldn't normally characterize the miser or monk we've described as happy. But the text suggests that Mill chooses (b) without admitting that he is therein abandoning the equation of happiness with pleasure:

"What for example, shall we say of the love of money? There is nothing originally more desirable about money than about any heap of glittering pebbles. Its worth is solely that of the things which it will buy; the desires for other things than itself, which it is a means for gratifying. Yet the love of money is not only one of the strongest moving forces of human life, but money is, in many cases, desired in and for itself; the desire to possess it is often stronger than the desire to use it, and goes on increasing when all the desires which point to ends beyond it, to be compassed by it, are falling off. It may then, be said truly that money is desired not for the sake of an end, but as part of the end. . .

What was once desired as an instrument for the attainment of happiness has come to be desired for I its own sake. In being desired for its own sake it is, however, desired as *part* of happiness."

Primary Questions: Does the miser/monk want something other than happiness for its own sake? Is this the result of some form of ignorance or error on his or her part?

Secondary Questions: Direct utilitarianism is the view that we *ought to* maximize happiness. Indirect utilitarianism is the view that we ought to adopt those rules that would best promote happiness. But Mill sets out to show that happiness is the most *desirable* thing and the only thing desirable in itself or apart from its effects. What is the missing step? Answer: that we ought to promote (or even maximize) that which is most desirable. Can we give an argument to support this missing step? Does it need an argument in its support? (If you don't think this assumption requires support, wait until we read Kant.)

Four questions:

(1) Even if we grant Mill the premise that we ought to promote what is most desirable, has Mill shown that we should (a) *maximize* happiness or only that we should (b) *promote some* happiness? Could a rational egoist (or someone who defends the rationality of acting selfishly) use this argument to support his own position? Why should we value the perspectives of other people we don't already love or care about?

(2) Is the notion of happiness Mill uses in this argument the same notion he uses in explaining the principle of utility? What's the relationship between pleasure and the heterogeneous collection of activities and states of affairs that one desires for themselves irrespective of their consequences? If Mill uses "happiness" in two distinct senses, does this present a problem for his "proof" of the principle of utility?

(3) Does the fact that happiness is desirable follow from the fact that everyone desires happiness? (See Mill's comparison of his project with proving that something is visible by showing that people see it.) Does 'x is desirable' mean x ***is*** *desired* or, instead, *that it is appropriate or good that x be desired*? Again, what is the connection between desiring something about which one has relatively full and accurate information and that thing's being desirable? Might S be wrong to desire X even though S is not mistaken about X's more or less value-neutral properties?

(4) Might a miser desire money for its own sake even when it is obvious to him or her that money is not bringing him or her pleasure? Suppose this is possible: One might think the miser is wrong to value money for its own sake, as money is really just valuable as a tool for the acquisition of things of independent value. Is this right? Does the miser make a mistake in collecting money instead of using it to purchase

pleasurable objects and activities? Does Mill think the miser is mistaken, or does he think that the miser's non-instrumental desire for money entails that the miser's happiness consists in the acquisition and retention of currency? If the miser makes a mistake in valuing money for its own sake, is the person who desires to be virtuous independently of his desire to promote happiness through benevolence guilty of a similar mistake? Is it a mistake to desire virtue for its own sake?

VII. Mill on Justice and Utility: The Normative Consequences of Psychological Accounts of Our Moral Judgments

"In the case of this, as of our other moral sentiments, there is no necessary connection between the question of its origin, and that of its binding force" (p. 87).

Mill's Distinction Between Natural Instincts and Reliable Instincts: It is not the case that if we have an innate or universal propensity to feel or judge certain kinds of acts to be unjust, then these acts really are unjust. Innate and universal propensities or instincts can be unreliable.

Mill's argument for the claim: Some of our instincts to act are bad: anger can lead to actions we judge imprudent when thinking back on them in a "cool hour," and appetites can lead to behavior we judge gluttonous or unhealthy upon reflection. Similarly, illusory experiences can lead to judgments we come to recognize are false, and there may be universal or innate forms of inference we judge to be fallacious upon reflection (e.g. the gambler's fallacy).

Indeed, we need to be especially careful in these cases, because when we are naturally or instinctively disposed to make certain judgments, we typically assume that these judgments must be true or reliable. But this assumption may not withstand scrutiny.

So we can ask:

(1) Regarding Mill's "Proof" of the Principle of Utility: We assume that a person's happiness is composed of those experiences she continues to want (for their own sake) even after she has reflected on them in a clear-headed, informed manner. We are disposed to think that what we desire (in these conditions) really is desirable or good for us. But might we discover upon further reflection that even our considered desires provide us with a largely **unreliable** guide as to what is truly good for us?

(2) We assume that those things that intuitively or immediately strike us as unfair or unjust or strike us as a violation of a person's moral rights really are unfair, or unjust, or a violation of a person's rights. But might we discover upon reflection that our intuitive or instinctual judgments regarding justice and injustice provide us with a largely **unreliable** guide as to what is truly just and unjust?

Question: Mill seems to dismiss without argument the possibility that what is good for us is something entirely different from what we want upon reflection in light of accurate information about it. But if he does dismiss this possibility out of hand, how can he consistently think that the Kantian (or deontologist) must seriously consider the possibility that our intuitive or instinctive judgments about injustice are largely unreliable? Note though that Mill does not reject judgments of justice altogether. Instead, he thinks the is a moral core to our sense of justice and that it is derived from considerations of happiness or utility in an indirect way.

VIII. Mill's Attempts to Characterize our Intuitive or Ordinary Judgments of Justice

A. Violation of Legal Rights: We call "unjust" the deprivation of property or liberty protected by the laws of the land. Things are not clear when we judge that the law is a bad one and the property or liberty denied someone is property or liberty they ought not to have been given. (E.g. protecting an escaped slave.)

B. Bad Laws: We call unjust the enacting of laws that violate a person's moral rights: laws denying racial, ethnic or religious groups the rights to vote, speak, move freely, etc.

C. Desert: We call it just when bad things happen to people who intentionally do bad things (so long as the bad received is proportionate with the bad paid out). We call it "unjust" when bad things are done to good people or when bad people get away with doing bad things without adverse consequences.

D. Violations of Contracts (Both Explicit and Implicit): We call it unjust when someone breaks a promise or violates what we take to be the legitimate expectations of others unless this is done to respect an obligation we judge weightier.

E. Partiality: We judge it unjust when officials or those entrusted by a group with representing its interests or distributing goods or services on its behalf fail to act in a wholly impartial manner by favoring their own interests or the interests of their friends and families over the interests of strangers. (E.g. judges must deliver their verdicts in an impartial manner, teachers must grade impartially, government officials must award contracts in an impartial manner etc.)

F. Equality: It is said to be unjust that a few should have privileges and rights denied to the many, though (Mill says) the interpretation of this claim is mired in controversy.

Question: Surely, Mill should conclude that the word "justice" is ambiguous or that its meaning is in some less drastic way sensitive to features of the context in which it is used. Why then does he insist that it is "not yet regarded as ambiguous"? Is he being sarcastic?

IX. The Genesis of Judgments of Justice

We need to look at the judgments of justice we make in each one of these areas and ask: (a) the psychological question: what causes us to judge acts unjust in these cases? And (b) the normative question: once we reflect on the sources of our judgments of justice do we find ourselves questioning the reliability or truth of those judgments or does our confidence in these judgments remain in place?

Mill's Psychology of Judgments of Morality: The idea of a law and conformity to law is the "primitive element" in our idea of morality. In the paradigm case we judge an act unjust because we have determined that it runs contrary to the law. With two exceptions: (a) We can think about the laws themselves and judge them to be unjust. (b) There are unjust acts we don't think should be prohibited by laws because of the negative consequences of enforcement. (As already discussed in class, we don't want the cops trying to figure out which girls are cheating on their boyfriends.)

How Mill defines "immoral" (or "morally wrong") within his utilitarian framework: "We do not call anything wrong, unless we mean to imply that a person ought to be punished in some way or other for doing it; if not by law, by the opinions of his fellow creatures; if not by opinion, by the reproaches of his own conscience. This seems the real turning point of the distinction between morality and simple expediency."

Notice that Mill here explicitly considers utilitarian views of the sort Singer endorses according to which we are "obligated" to do everything we can to promote utility and that he fails to endorse these views. This provides additional evidence against interpreting Mill as an advocate for direct utilitarianism or effective altruism.

> No one has a moral right to our generosity or beneficence, because we are not morally bound to practise those virtues towards any given individual. And it will be found with respect to this, as to every correct definition, that the instances which seem to conflict with it are those which most confirm it. For if a moralist attempts, as some have done, to make out that mankind generally, though not any given individual, have a right to all the good we can do them, he at once, by that thesis, includes generosity and beneficence within the category of justice. He is obliged to say, that our utmost exertions are due to our fellow creatures, thus assimilating them to a debt; or that nothing less can be a sufficient return for what society does for us, thus classing the case as one of gratitude; both of which are acknowledged cases of justice. Wherever there is right, the case is one

> of justice, and not of the virtue of beneficence: and whoever does not place the distinction between justice and morality in general, where we have now placed it, will be found to make no distinction between them at all, but to merge all morality in justice. (p. 95)

Mill never says that he **agrees** that we should "merge all morality in justice" and he seems to **distance himself** from those who argue that "mankind generally…have a right to all the good we can do them." Instead, he seems to embrace the view of obligation sketched above:

> **Mill's account of immorality**: an action is (pro tanto) morally wrong or impermissible if it violates the kind of social rule that should be enforced with social or legal punishments, where a social rule should be enforced with punishments if doing so promotes utility on the whole.

<u>Questions</u>: Is this a good definition of immorality? Might someone act immorally even if there is nothing to be gained from punishing actions of the sort she has engaged in?

Mill recognizes that his utilitarian account of immorality or the violation of rights does not entirely mesh with his audience's sense of justice. As we will see when we read Kant, there are many people who believe in poetic justice and think people should suffer for their misdeeds even when this is not necessary to promote happiness or mitigate suffering.

> I conceive [i.e. admit] that the sentiment [of justice] itself does not arise from anything which would commonly, or correctly, be termed an idea of expediency [here understand as the promotion of utility]; but that though the sentiment does not, whatever is moral in it does. (p. 95).

He goes on to analyze this (non-utilitarian) sentiment (possessed by his opponents and all those who do not already think of rights and obligations in a utilitarian way) into two essential ingredients:

> The two essential ingredients in the sentiment of justice are, the desire to punish a person who has done harm and the knowledge or belief that there is some definite individual or individuals to whom the harm has been done" (p. 95).

How did these "ingredients" arise? Mill gives a evolutionary explanation of how people come to desire the suffering of others. The desire to repel attacks is said to be universal and probably innate as it is shared by the primate species from which we've evolved. Mill says the same thing about the desire to retaliate: it is probably innate as it's inherited from our evolutionary ancestors. (Indeed, Karen Wynn and Paul Bloom at Yale have recently published experiments that show infants as young as 3-months-old like it when hinderers are punished and helpers are rewarded.)

There are of course differences between our desires for punishment and the desires of non-human primates. But according to Mill, the desire for punishment enjoyed by human beings primarily differs from the similar desires of non-human animals in that humans are capable of **greater sympathy** and have more **general intelligence**. Because of our greater sympathy we often want people to suffer for harming those who are unrelated to us (neither friends nor family). And because of our greater intelligence we can think of large groups of people (nations, etc), assign interests to these groups (by thinking about what's "good for America" etc.) and (by sympathizing with the group as a whole) we can come to desire the suffering of those who intentionally frustrate the group's interests and the interests of members of the group we've never met (e.g. outrage at attacks on "us").

But recall that Mill wants to do more than describe how we come to have the sense of justice we in fact have. He has, instead, a critical aim: to derive the conclusion that judgments of justice are only reliable when they are grounded in considerations of utility. "I conceive that the sentiment [of justice] does not arise from anything which would commonly, or correctly, be termed an idea of expediency; but that though, the sentiment does not, whatever is moral in it does" (p. 95).

"To have a right, then, is, I conceive, to have something which society ought to defend me in the possession of. If the objector goes on to ask why it ought, I can give him no other reason than general utility. If that

expression does not seem to convey a sufficient feeling of the strength of the obligation, nor to account for the peculiar energy of the feeling, it is because there goes to the composition of the sentiment, not a rational only but an animal element, the thirst for retaliation; and this thirst derives its intensity, as well as its moral justification, from the extraordinarily important and impressive kind of utility which is concerned" (p. 98).

Question: If we accept Mill's account of the genesis of our (non-utilitarian) intuitions about justice and rights does that undermine or "debunk" those intuitions that cannot be defended with considerations of utility?

Criticism: Mill hasn't shown that the only part of our judgments of justice worth retaining are those that advance overall utility. Mill might argue that **believing that some X ought to suffer** for harming someone or violating her rights is nothing beyond **wanting X to suffer for the harm she's caused**. But it's hard to see how any such argument could succeed. For instance, I can believe that a family member ought to suffer for harming others and yet find myself unable to desire his suffering. I can know that my spouse has acted unjustly in committing a string of robberies even as I do everything I can to help her escape the authorities.

Question: Mightn't we retain sophisticated judgments of justice which are abstracted from the thirst for revenge in this way, but which are not justified by considerations of utility? Consider in this light the analysis of our sense of justice defended by Rawls and the heirs to his (Kant-inspired) form of political liberalism.

Conclusion: There remains to this day a debate between utilitarians and deontologists about the relationship between judgments of justice and judgments of utility. And utilitarians (e.g. the neuroscientist Josh Greene) are still trying to "debunk" those of our intuitions about justice that do not conform to the utilitarian conception of morality. We now turn to a deontic conception of this kind with an examination of Kant's *Groundwork*.