

**AC 2008-1159: ETHICAL THEORY FOR ENGINEERS: AVOIDING CARICATURE  
AND INFORMING INTUITIONS**

**Travis Rieder, University of South Carolina**

## **Ethical Theory for Engineers: Avoiding Caricature and Informing Intuition<sup>1</sup>**

A professor of engineering ethics recently commented that she finds it difficult to teach ethical theory to engineers, as many students are strongly attracted to an intuitionist model of ethics – intuitionist here meaning that the students tend to make ethical judgments based largely on how a situation strikes their ‘moral sensibilities.’ One reason for this may be that engineering students are often, by virtue of their chosen discipline, not accustomed to the kind of abstract, “ivory-tower” thinking in which theoretical ethicists engage. In addition, teachers of engineering ethics tend to focus on case studies with very little emphasis on theory. Thus, it is understandable that engineering students might be more disposed to think carefully about individual cases and rely on felt intuition, rather than to endorse a theoretical ethical model and apply it to various cases. Thus, if there is any truth to the observation that engineers are often attracted to intuitionism in ethics, it is worth thinking carefully about what we, as teachers of future engineers, want our students to take away from an engineering ethics course, and what way of teaching ethical theory may best achieve our goals.

I take up this challenge by criticizing the way in which ethical theory is often taught in case study based courses, and then proposing a new model. However, I do not here question the general framework of utilizing both ethical theory and case studies. I am open to the possibility that such a criticism is needed, and that both the model I am criticizing and my proposed replacement are inferior to a radically revisionist way of doing engineering ethics; however, I am not qualified to take up such a study. I am not a specialist in engineering ethicist, with some vision of how the discipline ought to be organized; rather, I am an ethical theorist, who was invited into a conversation involving how best to make theory understandable to students of engineering. For this reason, my conclusion will apply only to those who believe that ethical theory has a place in engineering ethics, and that its role is to inform a case study approach.

This paper is organized in the following way: in section I, I describe a popular style of teaching engineering ethics and identify what I take to be its core faults. In section II, I suggest a way to take the role of intuition seriously by making a discussion of ethical theory subservient to observations of widely-held intuitions. And finally, in section III, I turn the observations from the previous two sections into what I hope is a workable model for teaching engineers.

### **I. A Problematic Model**

Those who teach engineering ethics<sup>2</sup> are likely familiar with the following model: a short period of a semester is used to provide a sketch of what are taken to be the three dominant approaches to ethical theory: virtue ethics, utilitarianism and deontology. After this task is dispensed with, the class is free to move on to what both students and teacher are much more interested in: case studies. Given such a focus on reacting to particular examples and a lack of emphasis on theory, there is little reason for a student to feel compelled to make consistent her judgments in various cases, or to rely on a background

model of reasoning which will aid in judging difficult cases; thus, we have a model which trains students to react to ethical dilemmas based on felt intuition.

The logical next question, however, is, “why should we take such a focus on ethical intuitions to be problematic?” After all, we cannot, realistically, train young engineers to be full-blown ethical theorists, carefully working out and adhering to a coherent and consistent ethical system. Thus, we must accept that engineers are likely to rely on how various situations strike their moral sensibilities, and perhaps the best that can be hoped for from an engineering ethics course is that it provides opportunity for practicing ethical reasoning. While much in this reaction is likely correct, I still find the model described above problematic, for at least two reasons.

### ***The Problem of Caricature***

The first point that I want to make is that a quick look at the “three dominant branches of ethical theory” is almost always a caricature. The typical reading list includes primary or secondary source material from Aristotle for virtue ethics, Jeremy Bentham or J. S. Mill for utilitarianism, and Immanuel Kant for deontology. However, these thinkers are quite complex in their thinking, with the result that what is able to be summarized in one or two classes is almost certainly misleading. In addition, it is difficult to understand the motivations of such historically removed philosophers, with the result that one often takes their views to be ridiculous and easily dismissed. For instance, Aristotle infamously held that women were merely malformed men, thereby legitimating their second-rate citizenship.<sup>3</sup> Bentham thought that there was no difference between the pleasure derived from reading poetry and playing a mindless children’s game.<sup>4</sup> And Kant held that all instances of lying are wrong, claiming that one ought not to lie even to a murderer who is looking for his next victim.<sup>5</sup> I have often seen in my students an immediate dismissal of a theory based on these individual or cultural quirks, the result of which is a failure to take theory seriously.

One might legitimately respond to the above point by saying, “well, it may have the effect described, but each of the thinkers named did, in fact, believe the theses attributed to them. So what is problematic about teaching them?” The problem, as I see it, is that teaching a quick, easily defeated version of each major theory does no justice to the central insights captured by each. Although each major, historical development in ethical theory has been shown to have counter-intuitive implications, each theory began from at least one strongly compelling intuition. By quickly glossing great thinkers and revealing what most of us take to be devastating consequences of their arguments, teachers may unintentionally bias students against any attempt to make coherent their intuitions.

### ***The Role of Intuition***

In addition, teaching in the way described above leaves ambiguous, and perhaps even hides, the role of intuition in ethics. To be clear: such a claim is not uncontroversial, and there is no unanimously-held view concerning the *proper* role of intuition in ethics.

However, the position I will advocate takes its lineage from a strongly compelling thesis of John Rawls',<sup>6</sup> which although not unanimously-held, is well-respected.

The position I am referring to is called 'reflective equilibrium' and was coined by Rawls, but has been employed by many theorists since. The basic idea is this: ethics consists in a dynamic process of intuition informing theory, which then can serve to correct or make consistent intuition. The process is continuous and loops back upon itself – one can always legitimately ask whether a theory ought to be revised in light of conflicting intuitions and vice-versa. While Rawls took this to be a substantive advance over theory deadlock – that is, it allows us to do moral theory without solving notoriously difficult metaethical questions<sup>7</sup> – we, as teachers, can employ a weaker principle of reflective equilibrium. This weaker principle simply acknowledges that ethical theory is and has been informed by central, widely-shared intuitions. The three major ethical theories were not devised out of thin air; they began from foundational beliefs of their creators. If Kant had not believed that ethical claims were objectively true or false for all human beings in all situations, he would not have come up with the system he did. Thus, as we will see below, Kantian deontology grasps the central insight that many share concerning the *categorical* nature of ethical commands.

Thus one can see that, in addition to doing a disservice to the historical theoretical positions by caricaturizing them, it is easy for teachers of engineering ethics to fail to point out what was accomplished in the form of embodying intuitions. The great thinkers in ethical theory are largely famous for what they did in the way of taking an insight that most of us unreflectively recognize, acknowledging its importance, and making room for it within a system. If we were able to make this notion clear to our students, then perhaps we could also make clear to them some of the problems that their intuitions have. If we could make them understand why Kant thought he was committed to telling the murderer where his neighbor was, then perhaps we could make them more aware of the dangers of their own, presumably 'self-evident' intuitions.

## **II. Taking Moral Intuitions Seriously**

So how might one 'take moral intuitions seriously' in the way being advocated? My suggestion is this: when teaching ethical theory to engineers, perhaps lecture ought to be organized so as to tell a story about conflicting and evolving intuitions. In order to take our moral intuitions seriously and simultaneously to avoid giving a caricatured picture of historical ethical theory, one could give an overview of ethics as a conversation taking place between real but conflicting insights. In this way, *a discussion of ethical theory becomes secondary to a discussion of intuition*, reversing the tradition role between the two.<sup>8</sup> The ultimate goal of such a process would be to show the strengths of each position by revealing which insights it captures, but also to show what those insights, when taken to their logical conclusions, entail. I will here briefly comment on what I take each of these central insights to be.

### ***Virtue Ethics***

Virtue models claim that ethics is not about *doing* what is right and *avoiding* what is wrong; but rather, it is about *being* good and *training against* habits which are bad. Virtue ethicists typically believe that the scope of ethics extends to all behavior, as one can only be judged on *how they live*, not based on specific actions.<sup>9</sup>

The central insight in this model of ethics is not difficult to see, as it claims that questions of good and bad are about how to live, which is based on intimate information not always available to others. For instance, common among contemporary ethicists are wild thought-experiments involving trolleys running over innocent citizens with decisions to be made about whom to run over,<sup>10</sup> or individuals faced with killing one innocent person in order to save nineteen others from being killed.<sup>11</sup> Virtue ethics, however, questions the value of such experiments, claiming that what is much more important than a focus on rare possibilities is a focus on the steady, underlying self.<sup>12</sup> This is intuitive for most of us, as a good person may occasionally unknowingly do something wrong, or a thoroughly corrupt, evil person may convince all others of his righteousness by doing many right, public actions. The virtue ethicist changes the focus from those individual instances, which yield very little information, and commands instead that each of us live virtuously.<sup>13</sup> And when asked what it means to live virtuously, advice is typically given in the form of ‘*be a good person,*’ or more specifically, ‘*be a good engineer.*’ Essentially, do what the good person would do in a given situation.<sup>14</sup>

### ***Utilitarianism***

Utilitarianism is a maximizing version of consequentialism, which states that an action is right if and only if it yields the best consequences. ‘Best consequences’ is then cashed out as the outcome which produces the greatest balance of happiness over unhappiness. One can then define happiness in a variety of ways. For ease of explanation, I will here use the Benthamite ‘hedonic utilitarian’ formula, which claims that an action is right if and only if it produces the greatest balance of pleasure over pain.<sup>15</sup>

The central intuition captured by any consequentialist formula is that one ought always to bring about a better state of affairs rather than a worse one. The hedonic utilitarian then gives that intuition more content by claiming that one ought always to bring about the greatest balance of pleasure over pain. Not only does the theory have strong intuitive plausibility in its claim that we ought to bring about better consequences rather than worse, but it is also quite determinate, so long as one assumes that people are regularly able to calculate predicted utility. There is only one right answer, and finding it is simply a matter of calculation.

### ***Deontology***

Deontology is a class of views united by their rejection of consequences as determining the moral worth of an action. Kant’s famous ‘categorical imperative’ is often taken as the exemplar of a deontological view as it states that an action is wrong if and only if the principle on which one takes herself to be acting could not, without contradiction, be universalized.<sup>16</sup> Thus wrongness, for Kant, is not a function of bad consequences;

indeed, consequences play no direct role at all. Instead, wrongness is a function of an agent's being unable to consistently will both a principle of action and a world in which everyone follows that same principle of action. From such a categorical imperative we can derive principles such as "don't lie," which always hold, regardless of circumstances, because the act of lying always embodies a rational contradiction; in order to lie, one must simultaneously will the act of lying and a world in which everyone lies, which would undercut the institution of truth-telling, making lying impossible.<sup>17</sup>

Kantian deontology, then, provides us with principles of morality which embody the following intuition: some acts are simply wrong, by virtue of the fact that they are the kind of act which is wrong (not by virtue of contingent considerations such as consequences). In addition, if an act is wrong, it is wrong for anyone, at any time. Wrong action, according to deontology, is the kind of act which *no one* ought to do.<sup>18</sup> This kind of *universalizability* is an important characteristic of morality for many people.

### ***Conclusions***

According to the preceding arguments, most of us, much of the time, have the following three intuitions concerning ethics: (1) ethics is primarily about *being* a particular kind of person, and evaluations of acts is secondary; (2) an action is right in so far as it brings about a better state of affairs than any alternative; and (3) an action is right if it is the kind of action which would be right for anyone to do, regardless of contingencies such as consequences. However, as should be obvious, these intuitions have the potential to collide – a possibility which seems crucially important for ethical decision-makers to learn. Therefore, in the final section, I will outline a model of teaching ethical theory in which the student's learning of particular theories is *not* the goal; rather, within this model, the role of theory is made subservient to the role of the intuitions which the theory captures. In this way, I hope to suggest a way of teaching engineering ethics that incorporates ethical theory as a dynamic conversation between widely-held, but competing and often-conflicting intuitions.

### **III. An Alternative Model**

It is important to acknowledge that I am not an engineer, and therefore do not have any real thoughtful advice about which case studies are most important for use in class or how exactly to incorporate them. I am here much more interested in beginning a conversation with engineers about how to incorporate ethical theory, with the hope that others might contribute thoughts on the engineering side, including choice and use of case studies. Although most of this section will include suggestions on readings, I want to begin with a note on why theory is important in the first place. Students may be resistant to the idea of studying ethical theory rather than simply diving into a particular case. Therefore, I might begin an engineering ethics class with a brief lecture on the importance of theory. While I won't transcribe an entire lecture here, the main point I want to make is as follows.

One interesting aspect of persons is our ability to give and assess reasons for our actions.<sup>19</sup> When interacting with others, people have the tendency to justify their actions, and we find some justifications better than others. One of the goals of pursuing ethical theory is in an effort to make the reasons we offer in defense of ethical verdicts stronger. While it may be unclear exactly what the strongest ethical system would look like, we do generally recognize a few values of ethical systems: for instance, an ethical system is better the more consistent, coherent and intuitively plausible it is. Any system offered which utterly lacks one of these three characteristics is likely to be (legitimately, we feel) thrown out of contention. Thus, it is worth our time to think about how the reasons we offer fit into a system of reasons, and how well such a system conforms to these values. Such a pursuit, we might think, adds to our moral integrity as agents.

In addition, ethical theory is helpful when it comes to new and difficult situations. While we may be comfortable with our ethical intuitions in common, everyday situations, new or difficult scenarios are likely to leave us uncertain about what action to take. If I were to rely merely on how the particulars of a situation strike my moral sensibilities, and had never been pushed to make them coherent in the way suggested above, then I would have little guidance in a situation where I had conflicting or underdetermining intuitions. Ethical theory can provide such guidance. While I will not here argue that young engineers should be pushed to commit themselves to a particular theory, familiarity with various theories, all of which capture important insights, can help to guide action in the way desired.

One final prefatory comment is in order: philosophers are unlikely to be pleased with what I offer here. This is because we are often romantic in our desire to train students in lofty, abstract theorizing. However, I assume that the goal of an engineering ethics course is not to force a liberal education onto engineers, but to provide them with tools for thinking more clearly about ethical decisions related specifically to their field. If this assumption is true, then a lecture concerning the beauty of Kant's system is inappropriate for this audience; instead, I here offer the first half of a model for teaching with the goal of making ethical theory real for engineers. The other half will be added when someone with the appropriate knowledge of the field puts together the following conversation of intuitions and thought-experiments with related case studies.

### ***Being Good***

Although it reveals a bias on my part, I would organize the theory section of an engineering ethics course mainly around the debate between consequentialist and deontological intuitions. This is not because I think conversations concerning *being good* has little to offer; on the contrary, I find much contemporary discussion of virtue valuable. However, as mentioned earlier, those interested in virtue tend to offer ethical advice in the form of 'do what the good person would do,' which many contemporary ethicists find problematic. In commanding only that one live as the good person would – perhaps supplemented with a classic list of virtues such as charity, benevolence, courage, etc. – one removes the possibility of action-guidance. Most of us turn to ethical reflection when we have questions: should I lie? Is it okay to steal if I'm really hungry?

Ought I to tell my friend's secret? But guidance in the form of 'be honest,' 'be just,' or 'be loyal' do not provide answers. For it is precisely in deciding what honesty, justice and loyalty entail that provides one with action guidance. Thus, for all its intuitive plausibility, virtue ethics fails to provide usable advice when we need it.

Engineering ethics is an applied branch of philosophy, and is therefore likely well-served by continuing to focus on case studies. Such an approach demands answers to questions of the form, "what should I do in this situation?" rather than "what kind of person should I be?" However, this is not to say that a conversation about virtue deserves no place in the syllabus. It seems a good idea to spend some time at the beginning of the course talking about not only what it might mean to be a good person, but about what it might mean to be a *good engineer*. Virtue ethics is exceptionally well-suited to ask what it might mean to be good in a particular role, which is a nice opportunity to show how ethical considerations inform our actions on a daily basis.

A reading from Rosalind Hursthouse's *On Virtue Ethics*<sup>20</sup> can be helpful here, as it is a contemporary look at what it might mean to be interested in virtue. In addition, there is plenty of room for talk about what virtues might be specific for an engineer, a discussion that would be well-served by anecdotes from a professional engineer's work experiences, either in the form of readings, the instructor's personal experience or a guest lecture. The idea would be to implant the idea that, regardless of the heated debates ahead about particular actions, there is an underlying self to be concerned with as well. The Hursthouse reading can then also serve to summarize criticisms concerning a virtue account's inability to guide action.<sup>21</sup>

### ***Bringing About the Best***

Most professors of ethics have a preferred historical reading for utilitarianism – a section from Bentham<sup>22</sup> or Mill,<sup>23</sup> which I suggest be thrown out. Just as I abandoned Aristotle in favor of Hursthouse above, I contend that, unless part of the goal of a course is for engineers to know something about history, Peter Singer is a much more sensible reading for an engineering class. I would take the primary selection from his *Writings on an Ethical Life*, in which he summarizes the central intuitions which he finds utilitarianism to so accurately grasp.<sup>24</sup> Singer's material can also be used to illustrate just how radical utilitarianism can be when taken to its logical conclusion, either through his article on famine and affluence,<sup>25</sup> or perhaps through his unique approach to animal ethics and his famous accusation of 'specieism.'<sup>26</sup>

Singer's message is simple: most of us believe that pain is bad and pleasure is good, and that, if given the opportunity, it is better for us to bring about good states of affairs than to bring about bad. However, if this insight is systematized and taken to its logical conclusions, we may not like where we end up. First, one might take issue with how 'good' and 'bad' consequences are typically calculated. The utilitarian, for instance, believes that each person 'counts for one and none for more than one,' such that the pain and pleasure which I feel counts for no more nor less than any other.<sup>27</sup> This yields the counter-intuitive implication that my pain should count no more in my ethical

deliberation than a stranger's, or perhaps even a future person's. Leaving aside the issue of *my* pain, we might still have a problem with the idea that my spouse, partner, or child's pain should count for no more than, say, a random person in a country around the world.<sup>28</sup> But Singer shows us that making consistent our shared ethical intuitions concerning pain and pleasure may well dictate giving aid to starving people in third-world countries rather than feeding ourselves and family at our current level, as the total balance of pleasure over pain would thereby be increased.<sup>29</sup>

One might also say that when a maximizing intuition is formulated into utilitarianism it becomes counter-intuitive because of its *aggregative* quality – that is, the fact that pains and pleasures can be summed across people. From this point we can make the more general criticism that utilitarianism takes what seems to be a sensible theory of rationality for the individual and applies it to all of humanity, which causes it to fail to recognize the separateness of persons.<sup>30</sup> This position claims that any view which allows small pains or pleasures to be summed across a vast number of persons in order to outweigh the much greater pain or pleasure of an individual must be mistaken.

Many short readings may be chosen to illustrate the counter-intuitive implications of utilitarianism. A few paragraphs from John Rawls' text makes clear why one might think that maximizing rationalities fail to take seriously the separateness of persons.<sup>31</sup> In addition, concerning the problem of aggregation, one might use a short thought-experiment taken from Alastair Norcross called 'Life for Headaches.'<sup>32</sup> According to Norcross, if it were the case that a person had to choose between relieving a massive number of mild, short-term headaches and saving an innocent person's life, utilitarianism tells us that there is some large but finite number of headaches such that one ought to relieve the headaches instead of saving the life. In essence, due to the aggregative quality of utilitarianism, massive numbers of small harms are able to be summed such that they can outweigh the harm involved in the death of an innocent person. But most of us think that no number of headaches is worse than the death of an innocent person – a widely-shared intuition which utilitarianism seems unable to account for.

The same criticism can be formulated in many other ways, some more concrete, such as the following example taken from T.M. Scanlon:<sup>33</sup> imagine that Jones has been injured in a television station's transmitter room, and that he is currently stable but regularly receiving painful shocks. In order to rescue him, the station must stop signal transmission, which would interrupt the transmission of the World Cup game for 15 minutes. If the rescuers were to wait for the game to finish, Jones would remain in severe pain for an hour. Most of us believe that no amount of enjoyment from watching the soccer game could outweigh Jones' pain – the signal should be shut down and Jones rescued. However, utilitarianism – taking our intuitions about causing the 'best' scenario to its logical conclusion – states that there is some number of viewers large enough such that their displeasure at 15 minutes of interrupted service would outweigh Jones' severe pain for an hour, and the right decision would be to wait for the game's end. Such an example brings the point a little closer to home, although the best scenario would be for an engineer to add to this section a case study reflecting real-life difficulties – what sorts of ethical decisions might an engineer be forced to make which would involve the kind of

analysis above? It seems that examples should be easy to find: even the famous Challenger explosion, with the traditional interpretation of the manager's "amoral calculation"<sup>34</sup> to launch reflects not an amoral calculation at all, but a following of utilitarian reasoning.<sup>35</sup>

There are many other criticisms of utilitarianism, many of which are imaginative and fun to teach. My favorite of these include Derek Parfit's "repugnant conclusion,"<sup>36</sup> as well as Robert Nozick's "experience machine."<sup>37</sup> Either of these readings would be valuable additions to a syllabus like the one I am creating; the only criterion is to show that a systematizing of the very reasonable utilitarian intuition can have radically counter-intuitive implications. Those who believe that maximizing the good is *the* criterion of wrongness are committed to the following claim: there is no act so bad that I ought not to commit it, should its commission prevent more acts of the same kind from being committed by others. This is a thesis which many of us find questionable.

### ***Acts that are Just Plain Wrong***

Such a claim provides the perfect opportunity for the deontologist to jump in: most of us think that some acts are, in fact, so bad that we ought never to perform them, *no matter what*. Rape is a likely candidate for such an act, as is the torture of a child; many of us think that no matter of consequences could legitimate performing one of these actions. In addition, we tend to think that if an act is right, then anyone in a similar position ought to be able to do it. It seems to count against an action if I am allowed to do it, but others would not be. The ethical thinking of Kant and his followers provide systematic accounts of such intuitions. The problem here is that the key argument of Kant's is very difficult for students to understand – especially students who have no interest or background in philosophy. Thus, under my alternative model I would summarize Kant in class, focusing on the fact that he embedded his ethical view in human rationality, which he saw as a way to yield determinate answers which were not dependent on a situation's outcome. Thus, lying is wrong because it produces a contradiction of *reason*, not because it would bring about a bad state of affairs. While students tend to find this unsatisfying, I would explain it as an attempt to build a system around the intuitions described above – the sense that some acts ought never to be permitted, and that if they are, then they are permitted for everyone.

I find deontology to be more difficult to bring 'up to date' than utilitarianism, so I would likely choose two, central, non-consequentialist aspects to focus on: duty and rights. W. D. Ross provides a good look at the kinds of duties one often thinks a person to have, granting intuition a starring role.<sup>38</sup> For a thoroughly anti-utilitarian account of rights, I might use excerpts from Nozick's *Anarchy, State and Utopia*, in which he defends the inviolability of human rights, regardless of consequential analysis.<sup>39</sup> Although this text is primarily a political treatise, it is founded on the insight that some things ought never to be done to a person, even if the alternative is anarchy – a nice illustration both of the belief that some actions are wrong regardless of consequences, and that if an action is wrong, it is wrong for everyone (including the government).

Although deontological views provide nice alternatives to the counter-intuitive implications of utilitarianism, they are not free from such implications themselves. Consider the so-called ‘paradoxes of deontology.’ In order to endorse deontology, one must accept that we sometimes have a duty to make the world morally worse. For instance: if it were the case that 20 other people’s lives were at stake and could be saved only by killing one innocent person, then following a duty not to kill would be to allow 20 deaths instead of bringing about one. Bernard Williams’ well-known thought-experiment dubbed ‘Jim and the Indians’<sup>40</sup> would, I think, serve well here, as it raises exactly this issue: if we want to reject maximizing rationality, we must be willing to allow 20 people to die in order not to actively bring about a single death. Williams’ article is well-suited to this collision of intuitions, because he is attempting to argue against a pure utilitarianism while realizing the power of maximizing rationality. It is not clear to Williams that Jim would be correct to allow 20 deaths for the sake of not bringing about one death himself.

Finally, Derek Parfit claims against deontology that its rejection of agent-neutrality can lead to self-defeating conclusions. If, as many deontologists believe, we have special duties to friends and loved ones (a fact which, if true, would refute Singer-style utilitarianism), then we are also committed to sometimes bringing about a situation in which our loved ones are worse-off than they would be under an agent-neutral ethic. Parfit calls such cases ‘Each-We’ dilemmas, which is a class of cases including prisoner’s dilemmas and what he calls ‘parent dilemmas.’<sup>41</sup> Consider the following case: A and B are both parents of three children each, and all six children are drowning. Both A and B are faced with the decision to save either one of their own children or two of the other’s. Because they are unable to communicate, they fulfill their ‘special obligation’ to their own children, with the result that they each save one of their own children instead of two of the other’s. Thus, because each followed a non-utilitarian, duty-based view, each parent ends up with only one surviving child when they could have each had two. In this version of ‘tragedy of the commons,’ focusing on obligations relative to oneself rather than agent-neutral, maximizing obligations, results in both parents faring worse in respect to keeping their children alive.

Here too, case studies from engineering would help to bring the point home. Such an example should not be hard to find, as Bernard Williams offers a thought-experiment which could easily be applicable to an engineer;<sup>42</sup> the basic idea of a problematic case of the kind in question is that sometimes doing one’s personal duty may result in a worse situation concerning the object of one’s duty. If an engineer who had strong personal feelings against war were to be faced with the prospect of designing weapons for the military, it may be that turning the job down out of a sense of duty could worse achieve one’s commitment to peace, as the position may be filled with someone who was excited and motivated to design weapons. Thus, one might find oneself back at the beginning intuition that, special obligations included, our ethical intuitions are better conformed to by utilitarianism than by any other theory.

### *Some Alternative Views*

Obviously, such an account is not the end of the story. The conversation concerning our widely-held ethical intuitions and how to accommodate them continues. Rather than leave students with the sense that all ethical theories are deficient, and that no matter how smart a theorist may be, some clever philosopher will invent a crazy thought-experiment which brings its counter-intuitive implications to the fore, I want to leave them with the picture that ethical theory is an ongoing conversation. Thus, before leaving the theory portion of the class, I would mention a few attempts to bring together these seemingly contradictory insights.

Although there is much interesting work being done in contemporary ethics which I think can be made available to non specialists, I will here just mention three. The first is rule-utilitarianism. Many contemporary thinkers find that a focus on rules instead of individual acts allows utilitarianism to account for both our maximizing intuition as well as strong, deontological intuitions concerning issues such as rights. Philip Pettit<sup>43</sup> and Brad Hooker<sup>44</sup> are well-known for advocating such a view, and short selections from either of both of their texts would be helpful here.

The second view is known as contractualism, the modern version of which claims its lineage from Rawls<sup>45</sup> but is advocated most strongly in its ethical form by T. M. Scanlon<sup>46</sup> and David Gauthier.<sup>47</sup> Contractualist accounts are bound together by the idea that some form of agreement, or hypothetical agreement, can ground ethics in a way that is intuitively plausible. Selections from an early essay by Scanlon would make a nice addition to the closing of this section.<sup>48</sup>

And finally, for an explicit look at how contemporary ethicists are attempting to account for the intuitions which have caused problems for other theories, Larry Temkin advocates a view which he calls, "Minimize Great Additional Burdens."<sup>49</sup> This view is essentially a reaction to the earlier criticism of utilitarianism, that it allows the small harms of many to aggregate and outweigh the large harm against one; in it, Temkin claims that a version of consequentialism is correct, but that instead of maximizing benefit, each person should attempt to minimize great additional burdens to individuals, and instead, to spread them over large numbers of people. Such an account gives us the intuitively correct answers to problems such as those raised by Norcross and Scanlon, without giving up the plausible focus on consequences.

It should be noted that many of the books and articles which I am citing as background for conversations of these intuitions are technical and difficult, and most ethicists would contend that to teach them to non-specialists would be foolish. However, the goal here would not be to teach the full articles, or to teach the arguments in depth; the goal is merely to capture the intuition raised by each example offered, a goal which I think is achievable. As an example, take my suggestion that the students read a piece from Scanlon's contractualist account as a way to suggest an ongoing conversation between the conflicting intuitions described above. While Scanlon's work is certainly not easy-going, a small piece or two could be excerpted, which provides students only with a picture of the central insight. So in the case of Scanlon, the take-home point is that a wrong action is one which could be reasonably objected to by others. Essentially,

Scanlon attempts to provide an alternative to utilitarian reasoning which remains intuitively plausible by grounding the criterion of wrongness in justifiability – an act is wrong if it cannot be justified to others. To me, such an account, which does not require a lot of difficult philosophical material, is quite plausible.

#### IV. Conclusion

One of the goals of the model of teaching described is to help future engineers realize that what they may take as sensible intuitions when analyzing a case study – intuitions such as, “It is never okay to risk human lives for the sake of fiscal savings – the Challenger should not have been launched when the engineers involved had doubts as to the safety of the conditions” – may get them into trouble when taken to their logical conclusions. The goal is to get the student to ask, “What kind of statement am I making when I say, ‘it’s never okay to do X’?” The realization that it is a deontological statement will hopefully better inform the conversation, by making the speaker aware of its implications. It may be that the student is okay with the realization that she is making a deontological claim – she may be prepared to defend against the utilitarian criticisms of her view. But the very realization of how one’s intuitions fit into a larger scheme, and the weaknesses of the scheme, seems to me to be a great advance over the situation in which we leave many of our students.

#### Notes:

---

<sup>1</sup> The existence of this project is due to a suggestion by Ann Johnson, for which I am grateful. I am also grateful to Justin Weinberg for reading multiple drafts, and for detailed comments and criticisms. Finally, for helpful conversation on the penultimate version of this paper, I want to thank the following individuals: Rich Holmes, Matt Kisner, Jonathan Krull, Scott Mayberry, Sonya Ozbey, Gordon Purves, and Chris Tollefsen.

<sup>2</sup> The following statement likely applies to many forms of professional ethics. Indeed – when a mandatory lecture on professional ethics was given to the research students at the University of South Carolina’s Medical School, I obtained a copy of the lecture notes, which was given by a professor of the Biomedical Sciences Department; the notes of the speaker’s lecture conformed exactly to the style I describe in the paragraphs following, indicating that this way of teaching ethics to non-specialists likely extends beyond engineering.

<sup>3</sup> Aristotle, *On the Generation of Animals*, II, III, 737a25.

<sup>4</sup> Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation*, ch. I.

<sup>5</sup> Immanuel Kant, “On a Supposed Right to Lie because of Philanthropic Concerns,” *Grounding for the Metaphysics of Morals: With On a Supposed Right to Lie Because of Philanthropic Concerns* (Indianapolis: Hackett, 1993), 63-68.

<sup>6</sup> See John Rawls’ *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971), 48-50. Although introduced by the name reflective equilibrium in this work, Rawls had devised the outline of the method in 1951 in his “Outline for a Decision Procedure for Ethics,” *The Philosophical Review*, Vol. 60, No. 2 (Apr., 1951), pp. 177-197.

<sup>7</sup> That is, we can move on to normative ethical theory without postulating an adequate solution to G. E. Moore’s ‘open-question argument,’ or without a solution to sticky questions concerning the metaphysical and epistemological status of moral facts. Reflective equilibrium, when introduced in its earliest form by Rawls in the 1950’s, allowed normative theorists to bypass metaethical gridlock which had largely killed ethical debate for the proceeding fifty years.

---

<sup>8</sup> In this paper's original formulation, I made the more modest suggestion that theory should be taught specifically with an eye to informing intuitions. The suggestion that my point is made clearer and more interesting by radicalizing it into theory's *subservience* to intuition is a suggestion that I owe to Justin Weinberg.

<sup>9</sup> A focus on virtue is common with various naturalist models of ethics, in which there is a given telos or end to human existence, with the good life being that life which most closely achieves its telos (See Aristotle's *Nicomachean Ethics*). It is also popular with various religious ethics, with a way of life being commanded by God in the form of particular virtues (chastity, benevolence, charity) (See St. Thomas Aquinas' *Summa Theologica*), as well as with a small, contemporary movement often called 'anti-theorists' (See Alasdair MacIntyre's *After Virtue*).

<sup>10</sup> Although quite famous in contemporary ethical debates, I believe this thought-experiment originated with Phillipa Foot in her "The Problem of Abortion and the Doctrine of Double Effect," *Oxford Review*, 5, 5-15.

<sup>11</sup> Again, many versions of such a thought-experiment exist, but the particular version I have in mind is often referred to as "Jim and the Indians," and is given by Bernard Williams in *Utilitarianism: For and Against*, edited with J.J.C. Smart (Cambridge: University Press, 1973), 93-100.

<sup>12</sup> For a clear, contemporary version of this argument, see Rosalind Hursthouse's *On Virtue Ethics* (Oxford: University Press, 1999).

<sup>13</sup> Such a conversation extends back as far as Plato, who, through the mouth of Socrates, continually asks the question, 'what is justice *in the soul*?' When one considers that a thoroughly just man may be tortured and killed, with his image soiled, while a thoroughly corrupt man may appear virtuous to all others and gain the rewards of virtue, while keeping his corruption to himself, then one must be able to give an account of justice as it appears only within one's soul. Such a description is the project of the bulk of the *Republic*.

<sup>14</sup> Hursthouse, *On Virtue Ethics*, 26-31.

<sup>15</sup> Bentham, *Introduction to Principles*, ch. 1.

<sup>16</sup> This is a broad summary of Kant's argument, found in the *Groundwork for the Metaphysics of Morals* (Cambridge: University Press, 1997), esp. pp. 31-34.

<sup>17</sup> This is often explained by using some sort of 'test of universalizability.' Consider the following explication: I am considering whether to lie to my friend, Alex. Kant tells me that such an action is right only if I can rationally, without contradiction, will both the action and a world in which everyone follows the maxim on which I take myself to be acting. In this instance, I may be following the maxim, "Lie to your friend if it is convenient." However, in a world in which everyone followed such a maxim, lying would become impossible, as lying requires a general institution of truth-telling – an institution which is undermined by everyone following my maxim. Thus, Kant tells me that I should not lie to Alex, as I cannot, without contradiction, will both that I lie and that everyone lies.

<sup>18</sup> That is, holding constant morally relevant features. It may be wrong for a stranger to treat my child in all sorts of ways that would not be similarly wrong for me. The central idea here is one that was captured by Kant when he claimed that immorality is about taking exception for oneself – acting on a rule that I cannot will that everyone act on (*Groundwork*, 33-34).

<sup>19</sup> This is the view which serves as a foundation for T. M. Scanlon's contractualist account of morality. See further Scanlon, *What We Owe to Each Other* (Cambridge: Harvard University Press), 1998, 105-106.

<sup>20</sup> Hursthouse, *On Virtue Ethics*. Although the exact selection could vary according to the instructor's goals, the most helpful sections would likely be found in the Introduction, pp. 1-16 and ch.1 on "Right Action," pp. 25-42.

<sup>21</sup> *Ibid.* Hursthouse gives a nice summary of the arguments against a virtue model's action-guiding ability on p. 26.

<sup>22</sup> Bentham, *Introduction to Principles*.

<sup>23</sup> John Stuart Mill, *Utilitarianism* (Indianapolis, IN: Hackett, 2001).

<sup>24</sup> Peter Singer, "Introduction," *Writings on an Ethical Life* (New York: HarperCollins, 2000), *xiii-xx*.

<sup>25</sup> Singer, "Famine and Affluence," *Writings*, 105-117.

<sup>26</sup> Singer, "All Animals Are Equal...", *Writings*, 28-46.

<sup>27</sup> This is made explicit by Singer in his summary of utilitarianism's central intuitions.

<sup>28</sup> This is typically formulated as the charge that utilitarianism is "too demanding" to be taken seriously. Ethicists believe that ought implies can, and if it is psychologically impossible for us to grant the same

---

weight to all others as we grant to ourselves and intimates, then the theory fails to provide usable action guidance.

<sup>29</sup> For a famous version of this style of argument, see Peter Singer's "Famine and Affluence."

<sup>30</sup> This point is made by John Rawls in his *Theory of Justice*, pp. 26-27.

<sup>31</sup> A good selection would be pp. 22-27 of *Theory*.

<sup>32</sup> Alastair Norcross, "Comparing Harms: Headaches and Human Lives," *Philosophy and Public Affairs* 26, no. 2 (Spring, 1997): 135-167. I should note that this is a long and complex article; however, the counter-example of 'Life for Headaches' is presented on the first page, and a very short reading could easily be chosen just to present this particular argument against utilitarianism.

<sup>33</sup> T.M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998), 235.

<sup>34</sup> Diane Vaughn, *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA* (Chicago: University Press, 1996).

<sup>35</sup> It should be noted that such an interpretation is no longer unanimously accepted. Vaughn, who coined the term 'amoral calculation' (previous footnote), argues persuasively that the well-accepted interpretation of the Challenger tragedy as a result of engineers being unwilling to strongly voice their opinions and the managers being unwilling to listen to the engineers is over-simplified and misleading. See both Vaughn's *The Challenger Launch Decision* and Lynch and Kline's "Engineering Practice and Engineering Ethics."

<sup>36</sup> Derek Parfit, *Reasons and Persons* (Oxford: University Press, 1984), 381-387. This thought experiment claims to show utilitarianism's commitment to bringing about a world in which billions of people enjoy lives barely worth living instead of a smaller, blessedly happy population.

<sup>37</sup> Robert Nozick, *Anarchy, State and Utopia* (Basic Books, 1974), 42-45. The intention of this famous thought-experiment is to show that none of us desire only pleasure and the absence of pain.

<sup>38</sup> W. D. Ross, *The Right and the Good* (Oxford: University Press, 1930). The most helpful selection would be taken from "What Makes Right Acts Right," pp. 16-41, especially the first several pages.

<sup>39</sup> Nozick, *Anarchy, State and Utopia*. The central deontological insight here can be grasped from a short selection from the Preface, pp. ix-xi and a part of the section on "Moral Side Constraints," pp. 26-33.

<sup>40</sup> Bernard Williams, *Utilitarianism: For and Against*, edited with J.J.C. Smart (Cambridge: University Press, 1973), 93-100.

<sup>41</sup> Parfit, *Reasons and Persons*, 95-98.

<sup>42</sup> Williams, *Utilitarianism*, 93-100.

<sup>43</sup> Philip Pettit, "The Consequentialist Can Recognize Rights," *The Philosophical Quarterly*, Vol. 38, No. 150 (Jan., 1988), pp. 42-55.

<sup>44</sup> Brad Hooker, *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality* (Oxford: University Press, 2003). Although his book is difficult, the introduction lays out the intuitive argument for rule-consequentialism nicely, and even a reading as short as the first three pages (sec. 1.1) would be useful.

<sup>45</sup> See Rawls' "Justice as Fairness," Ed. Cahn & Haber, *20<sup>th</sup> Century Ethical Theory* (Englewood Cliffs, NJ: Prentice Hall), 1995; *A Theory of Justice* (Cambridge, MA: Harvard University Press), 1971; *Political Liberalism* (New York: Columbia University Press), 1993.

<sup>46</sup> See Scanlon's "Contractualism and Utilitarianism," in Amartya Sen and Bernard Williams, eds., *Utilitarianism and Beyond* (Cambridge: University Press) 1982, 103-128; *What We Owe to Each Other* (Cambridge: Harvard University Press), 1998, 153.

<sup>47</sup> David Gauthier, "Why Contractarianism?" Ed. Peter Vallentyne, *Contractarianism and Rational Choice* (New York: Cambridge University Press, 1991).

<sup>48</sup> Scanlon, "Contractualism and Utilitarianism." A short selection of the four or five pages dealing specifically with contractualism and Scanlon's criterion of wrongness would likely suffice for present purposes.

<sup>49</sup> Larry Temkin, "A 'New' Principle of Aggregation," *Philosophical Issues* (15, Normativity, 2005), 218-234. Here again, this article is long and technical. However, the presentation of the positive view is largely at the beginning, and the course instructor could easily excise a manageable reading.