

Lec 5

Objections to utilitarianism

Example: Dark Patterns

Dark patterns - Manipulative UX (User Experience) design

- Considered unfair and deceptive by the Federal Trade Commission (FTC) in the US in Oct 2021, by the EU in March 2022
- In Canada, there is no regulation explicitly addressing dark patterns, but we're going in that direction. Always keep in mind that the existence of a **legal loophole** doesn't mean an action is ethical.
- Existing privacy regulation already has some teeth so there's no guarantee that you and/or your company wouldn't face legal charges.





Personal Information Protection and Electronic Documents Act (PIPEDA)

What there is in Canada

- The Personal Information Protection and Electronics Document ACT ([PIPEDA](#)) relates to personal information – developed in the 90s. Enforcement of the act by the Office of the Privacy Commissioner of Canada has guidelines for obtaining meaningful consent regarding the collection, use and disclosure of personal information
- Canada's Anti-Spam Legislation (CASL) prohibits false and misleading statements in electronic messages
- If Bill C-27 got passed, the new Consumer Protection Act will have explicit language
- Section 16 of Bill C-27: “[An organization must not obtain or attempt to obtain an individual’s consent by providing false or misleading information or using deceptive or misleading practices. Any consent obtained under those circumstances is invalid.](#)”
- [6th Jan 2025 - Bill C-27 was dropped due to prorogation \(interruption\) of Parliament](#)

Manipulating you to do what is best

Consent and avuncular ethics

- Utilitarianism is perhaps particularly apt to examine an issue like dark patterns as the theory is sensitive to context – rejects absolutism
- Kantian ethics, as we will see, would reject all manipulation regardless of intention, wouldn't consider context, and the outcome would be irrelevant
- Utilitarian ethics could possibly accept manipulation depending on the outcome
- Examples:
 - 1. Trying to get someone to see as “default” something that is good for them. Declining contributing to pension plan is clear and possible but hard work – avuncular ethics
 - 2. Visual design - If the user keeps clicking yes without reading, they get much needed insurance for a dangerous activity. Declining insurance requires spotting and checking a tiny box – more paternalistic and violates consent more

Revisions

Some typical issues

- No consideration of objections
- Summarizing utilitarianism down to one sentence or two (e.g. “it’s about aiming for the general good”) does almost nothing to demonstrate knowledge of the theory.
- Quotes on utilitarianism (Bentham, Mill) unaccompanied by an explanation in your own words typically do a very poor job of explaining the theory (see next slide).
- Complex calculations of utility (X causes pleasure, Y causes pain, Z would cause even more pleasure, ...) do little to show knowledge of the theory. This just shows your knowledge of basic math!
- *Critical* discussion of problematic features (such as the ones we cover today) does a lot more to show knowledge
- Engage with complex issues successfully
- The approach of avoiding mistakes by only dealing with the basics guarantees a low grade

Revisions - Misinterpreting act utilitarianism

- **The student writes: “We should maximize utility for the majority”. Is this correct?**
- No
- In 1776, in his A Fragment on Government, Jeremy Bentham invoked what he described as a “fundamental axiom, it is the **greatest happiness of the greatest number** that is the measure of right and wrong”
- 1. Clarification: utilitarianism in fact does **not** say we must serve the majority. It just so happens that in many cases (but not all) utility adds up when high numbers of people are found (e.g. kill 1 to save 19)
- 2. Also, utilitarianism doesn’t recommend uniform and wider distribution of utility so we can serve a “greater number”
- **What matters is only total aggregated utility. Choose that option no matter how many are helped, and how distribution looks**

Revisions - Misrepresenting act utilitarianism

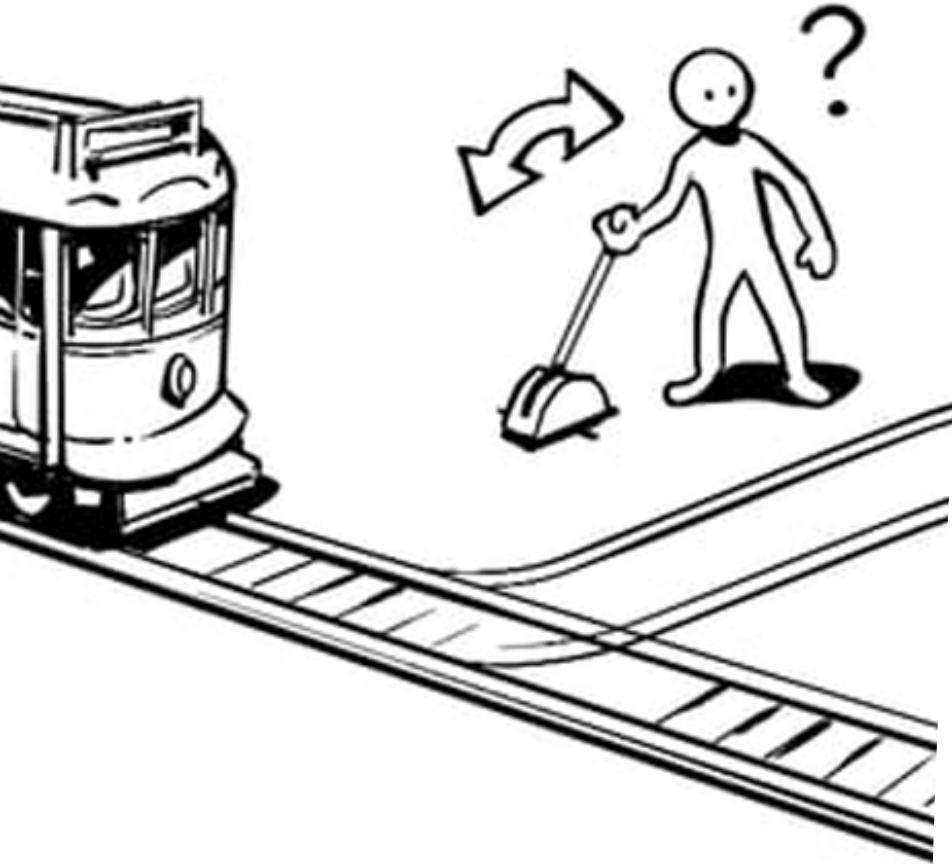
- **The student writes: “it’s moral because the benefits outweigh the costs”**
- This is an inaccurate summary of the actual procedure
- 1. You may have several options where the benefits outweigh the costs
- 2. You may have no options where this happens
- Student: You tell us not to use all these wrong definitions. What’s the right one??

Utilitarianism is not a cost-benefit analysis, which has a *point of view*. What costs and benefits **for X?**

Cost Benefit Analysis	
With Current Gas Tank	With Safety Alteration
180 burn deaths	Cost = \$11 per vehicle
180 serious burns	Total = \$137 million
2100 Pintos burned	
Costs = \$200 000 per death	
\$67 000 per serious injury	Second alternative = Rubber Bladder
\$700 per car	Cost = \$5.08 per vehicle
Total = \$49.5 million	Total ~= \$64 million

Utilitarianism is also not about causing more good than bad

Using/Applying better than any copy-paste definition



We never ask for super short single-sentence definitions. Utilitarianism, like any other fairly complex theory, cannot be summarized in a single sentence. Use the *whole essay* to show your knowledge.

The closest and **still incomplete** “short” definition will perhaps be to describe the “algorithm” and *implement* it:

- 1) identify all affected by available actions
- 2) for each available action, aggregate all relevant utility (since the theory subscribes to normative hedonism, this is the sum of pleasure and pain, using ECI) considering the *actual* outcome and
- 3) choose the option that maximizes utility

Revisions – Why is it called *ACT* utilitarianism?

- Suppose that someone says: “Respecting consent maximizes utility (as opposed to never respecting it), therefore we must always respect consent”. Is this what act utilitarianism says?
- No
- Clarification: *act (action)* utilitarianism evaluates actions. Respecting (all) forms of privacy is not an action (i.e. it’s a *class* of actions). Act utilitarianism is not absolutist and evaluates each action on its own merits
- This would be rule utilitarianism – choose the *rule* (not action) that maximizes utility
- Often accused of collapsing into act utilitarianism, as an even better rule for maximizing utility would be “respect consent unless in a specific case you can maximize utility through a breach of consent, such as using dark patterns”

Revisions - Misinterpreting act utilitarianism

- **The student says “The action is right because it causes happiness”. Is this right?**
- “Actions are right **in proportion** as they tend to promote happiness” – Proportionality doctrine of the Greatest Happiness Principle, Mill
- You may infer from the quote that a) right actions always cause happiness, b) there are many right actions and c) an action is right if it increases happiness. All incorrect.
- Clarification:
- A) The action recommended by utilitarianism may cause pain
- B) *Increasing utility* is not enough if it's not *maximized*. It's the “greatest” happiness principle.

First, the three objections to utilitarianism as discussed by Mill himself

1. Doctrine of swine

2. Too high for humanity

3. Lack of time

Objection to utilitarianism

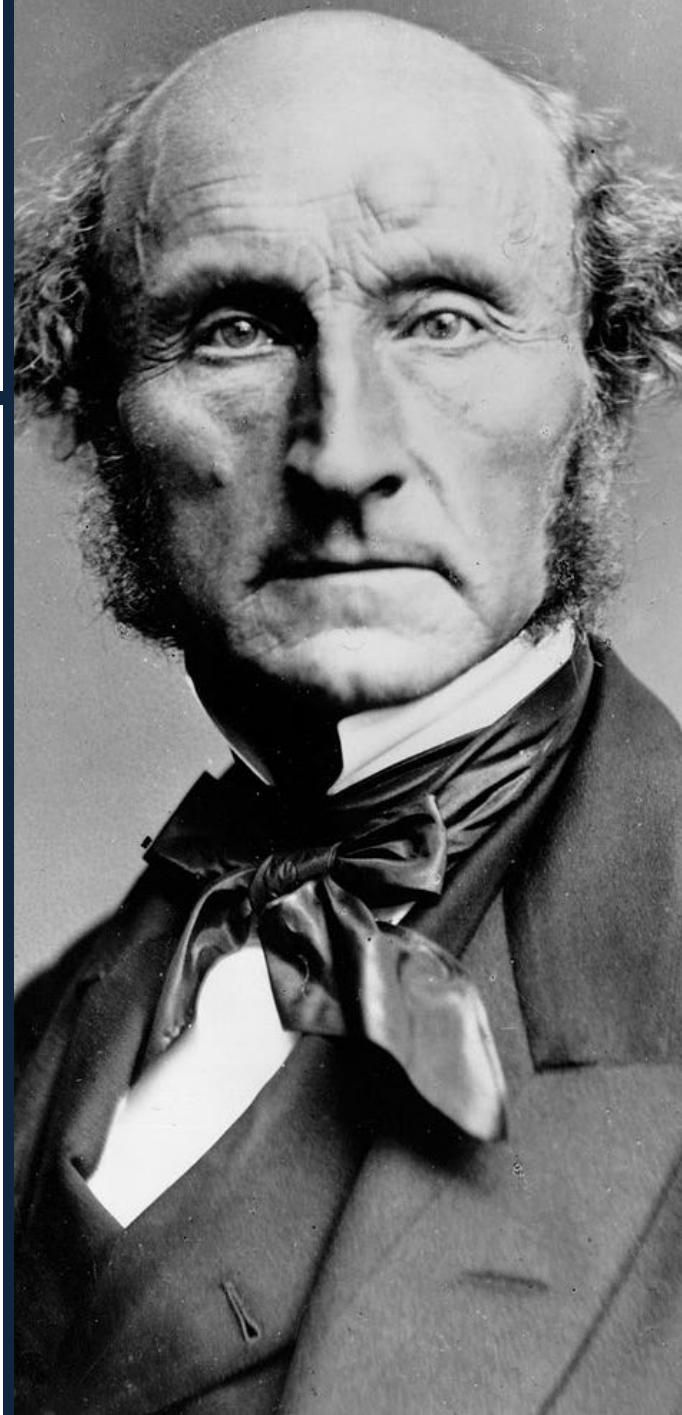
- Doctrine of swine

- “Pleasure and pain are our sovereign masters” -psychological hedonism
- We should maximize utility - normative hedonism
- Objection, as described by Mill: “To suppose that life has no higher end than pleasure – no better and nobler object of desire and pursuit – they designate as utterly mean and grovelling; as a doctrine worthy only of swine”
- The objection is based on utilitarianism’s claim that pleasure is the only thing that has intrinsic value (all other things have instrumental value).



Mill's response

- Mill: “some kinds of pleasure are more desirable and more valuable than others” – so we must consider quality as well as quantity
- How to know which are more desirable? Ask those who are properly acquainted with both, says Mill (the “competent judges”)
- **Is this a good response?**
- How do we know who is a competent judge (who is properly acquainted with the two pleasures to compare)?
- What if judges disagree?
- Doesn’t the idea of quality possibly smuggle non-utilitarian forms of morality evaluation to utilitarianism?



Objection to utilitarianism

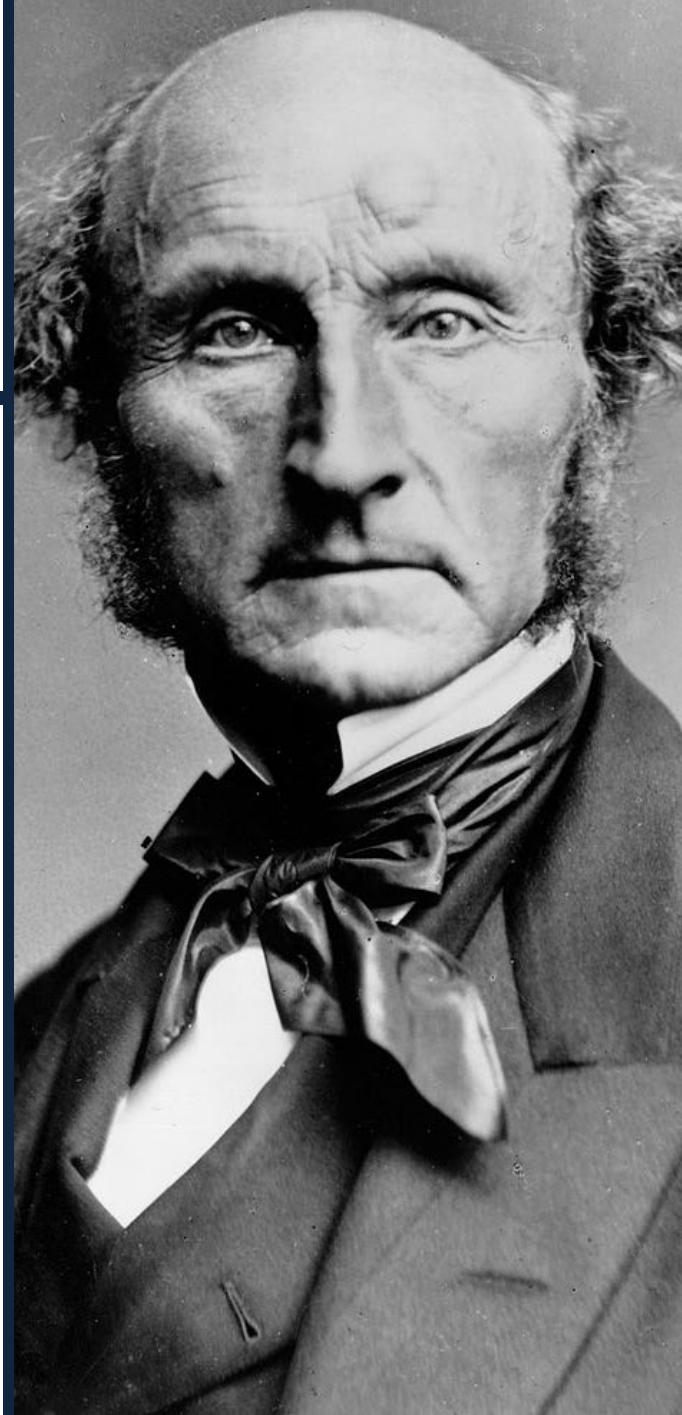
- Too high for humanity

- Mill's description of the objection - "It is exacting (demanding) too much to require that people shall always act from the *inducement* of promoting the general interests of society"
- Example of the objection: When we are kind to a neighbour, are we really **motivated** by the *general interest of society*? Surely, this cannot be the case. It is demanding too much.



Mill's response

- Mill says that this is confusing the “rule of action with the motive of it. It is the business of ethics to tell us what are our duties, or by what test we may know them, but **no system of ethics requires that the sole motive of all we do shall be a feeling of duty”**
- **But is it true that motives don't matter?**



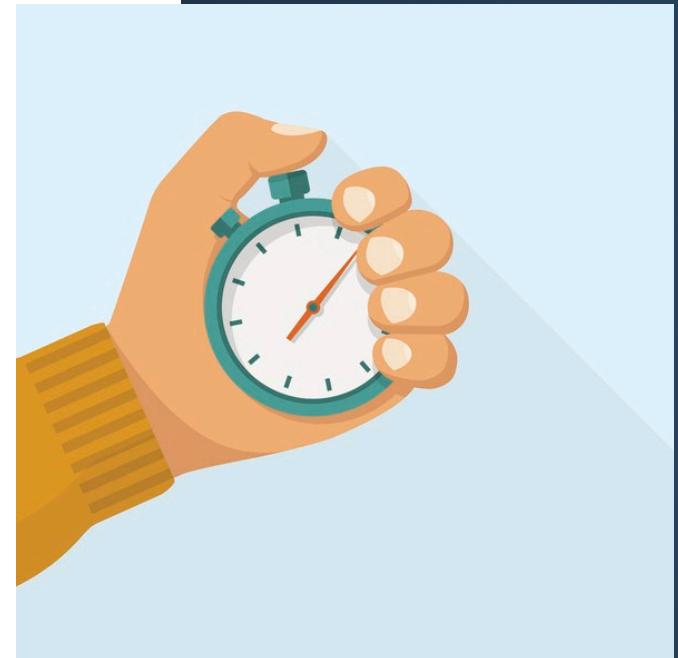


Is it true that motives don't matter in establishing right and wrong?

- Example from the Michael Sandel lecture: A doctor goes to a jungle to heal villagers but unknowingly she was carrying a disease and they all die
- Version B: A doctor goes to a jungle to wipe out all villagers. Knowingly she is carrying a disease she believes is lethal to them and they all die
- Version C: A doctor goes to a jungle to exterminate all villagers. Knowingly she carries a disease she believes is lethal to them, but they get immunity and are now healthier

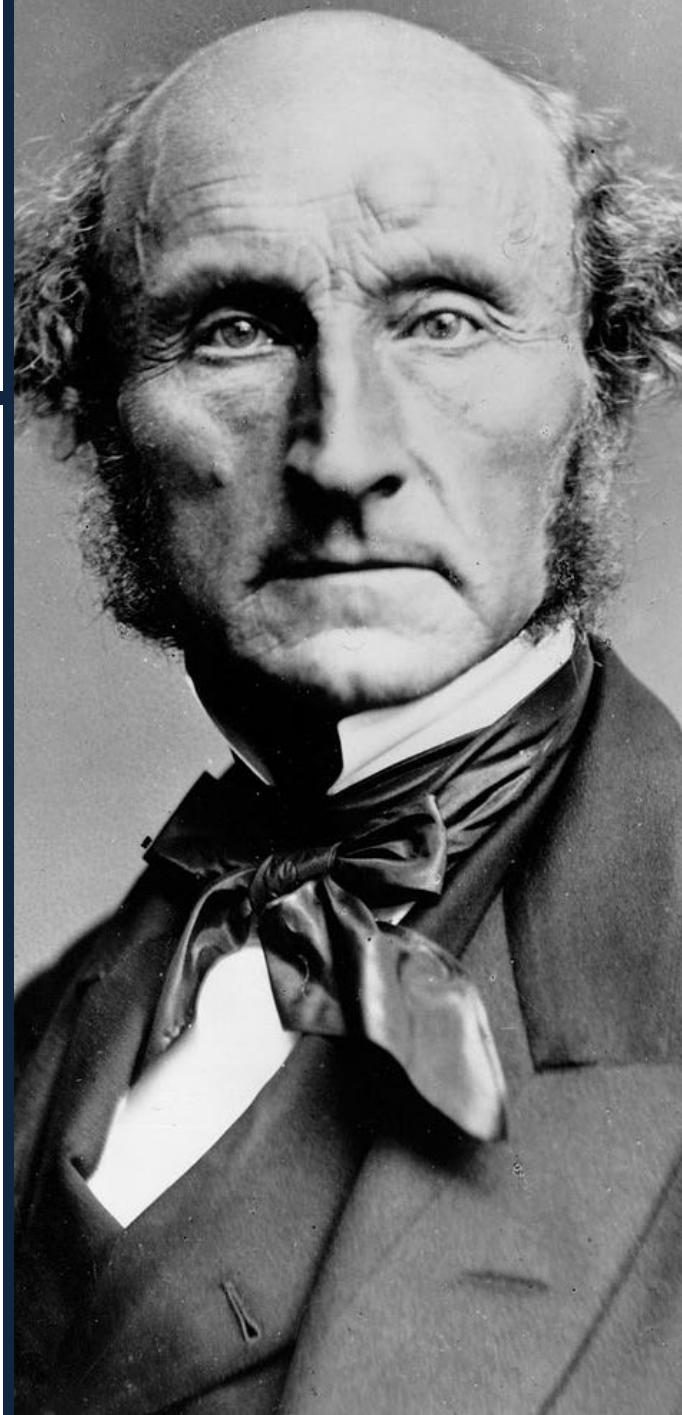
Objection to utilitarianism - Lack of time

- Mill's description of the objection:
"There is not time, previous to action,
for calculating and weighing the effects
of any line of conduct on the general
happiness"



Mill's response

- “There is no difficulty in proving any ethical standard whatever to work ill if we suppose universal idiocy to be conjoined with it; but on any hypothesis after that, mankind must by this time have acquired positive beliefs as to the effects of some actions on their happiness” Mill
- Translation: any normative ethical theory can be accused to fail to provide correct answers if we assume the person using it is an idiot. Let's not assume that.
- One can know, by experience, that some actions (e.g. lying, corruption, torture) are likely to cause pain, others to cause pleasure. There is no need to calculate every time. Bentham had already made that point.



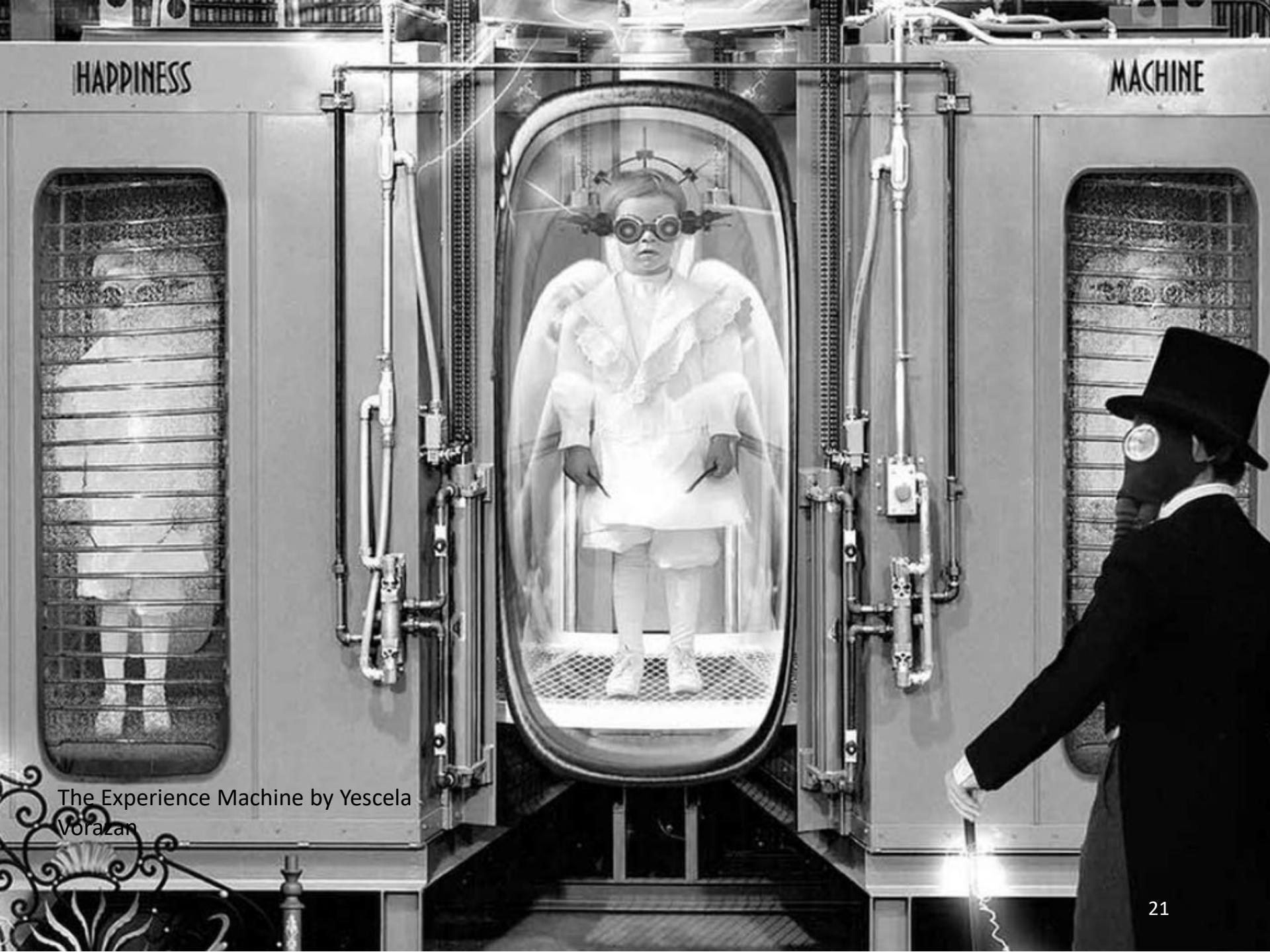
Additional problems with utilitarianism

- Motivational/psychological hedonism
- Normative hedonism
- Desire-fulfillment theory of welfare (Mill's proof)
- Aggregating utility
- Strong doctrine of negative responsibility
- Trivial actions
- Promises
- Punishment
- Justice
- Supererogatory action
- Impartiality

- Q: Is it really true that all our actions can be ultimately explained by reference to pleasure?
- A: Is my business/romantic partner cheating on me?
- Q: Is it really true that all our actions can be ultimately justified by reference to pleasure? (Is it true that we **should** act to maximize pleasure, because pleasure is the only good?)
- One famous argument/answer against both: Nozick's experience machine:
- If pleasure and pain are the *only* things that have ultimate value, you should want them even if they are not deserved/based on real accomplishments/etc.
- Marathon medal example



Psychological and Normative Hedonism



The Experience Machine by Yescela
Vorazan

Nozick's experience machine

- Anarchy, State and Utopia (1974)
- The experience machine:
 - 1. Can give you any experiences you want
 - 2. Perfectly simulates a life of achievement, pleasure, love, etc
 - 3. It is indistinguishable from reality
 - 4. You must plug in permanently (you won't remember choosing)
- **Do you plug in?**

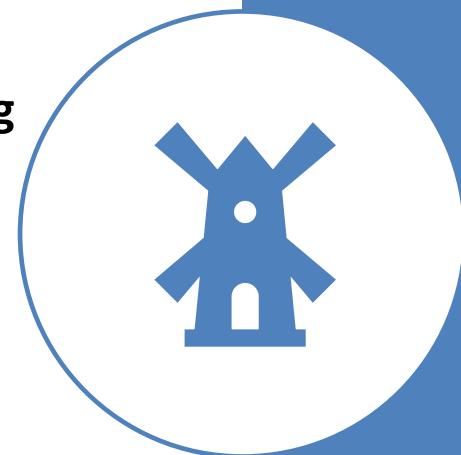


Nozick: 3 reasons not to plug in

- 1. *We want to do certain things, and not just have the experience of doing them.*
 - "It is only because we first want to do the actions that we want the experiences of doing them." (Nozick, 43)
(in other words, it's not the experience – the pleasure – that matters the most. Applying to our example, we want the person to be well, and secondarily to have pleasure *because they are doing well*)
- 2. *We want to be a certain sort of person.*
 - "Someone floating in a tank is an indeterminate blob." (Nozick, 43)
(in other words, character matters to us, rather than just pleasure – whether we are truly a good person, or just falsely believe we are a good person)
- 3. Plugging into an experience machine limits us to a man-made reality (it limits us to what we can make).
 - "There is no *actual* contact with any deeper reality, though the experience of it can be simulated." (Nozick, 43)
(in other words, it seems truth matters to us - **ignorance may be bliss, but it's worthless bliss**)
- Conclusion: pleasure is not the ultimate motivation + pleasure is not the only thing that has intrinsic value.

Desire-fulfillment theory of welfare (Mill's proof)

- Mill – We desire pleasure, therefore it is desirable
- 2 interpretations for “desirable”:
- We desire pleasure, therefore we desire pleasure – trivially true but circular
- We desire pleasure, therefore it is good – **clearly wrong**
- Mill argued each individual is better positioned than anyone else to know what is good for himself/herself. This is one argument he used to refuse interference from the state in personal matters that don't affect society.
- Objections:
- But don't humans desire all sorts of things that are bad for them?
- If I desire to be hurt simply because I have low self-esteem, is it desirable that I am hurt? No.
- If I desire crack because I'm addicted to crack, does that show that crack is desirable (good for me)? No.



Aggregating utility

Brief discussion

- **Is it OK to aggregate all utility? Or must we respect the rights of individuals, even at a cost to general happiness?**
- 1. The use of dark patterns leads a user to accept to share photos from his phone that he wished were private.
- But as it turns out in one particular case, eventually the RCMP uses these photos to identify him as a sex offender and get a conviction.
- **Was the original violation of privacy ethical?**



Strong doctrine of negative responsibility in Killing One to Save 19

- Bernard Williams asks whether we are really responsible for the things we *don't* do *as much* as we are responsible for the things that we do, asks Bernard Williams.
- Note that he's not denying that one can be guilty of negligence in certain contexts.
- Bernard Williams **integrity argument** – One may have vowed never to kill. Is not killing as blameworthy as killing? Perhaps an argument based on integrity has some purchase, thus placing the blame entirely or at least primarily on Pedro rather than oddly placing it on the person who refuses to hurt anyone out of moral integrity.
- Or imagine that after killing 1 to save 19, Pedro does release the 20, but then he gets another 20. Jim is now asked by psychopath Pedro to engage in increasingly horrible actions, again hurting 1 to save 19. The utilitarian will always make the same calculation, and will keep always doing what Pedro expects...

Trivial actions

- According to utilitarianism, utility must be maximized.
- This seems to entail that we can't take time out. Ever.
- Can't we just do something that is trivial, when we are not concerned with maximizing utility (e.g. buying new shoes, watch Netflix, go for a walk)? Do we *always* need to be maximizing utility?



How utilitarians see promises

- Promises will rarely be kept, since you only need to keep a promise when it maximizes utility
- When a promise is kept, there's no requirement to do it *because* it was promised.
- For utilitarianism, breaking a promise may cause **pain** (and that relevant for the theory), but otherwise there's no intrinsic moral obligation to keep promises



Example from the Sandel lecture: Promise to the dying man

- Man and grandfather go sailing.
- Shipwrecked on an island, the grandfather is dying and asks to be buried and that his body is treated with respect. The grandson promises but later considers whether to use the body for fishing bait.



Utility		
• Use body for bait		+25
• Bury body, according to promise	-10	
• Let body rot where it is		-2
• Throw body in sea		-1

- Discussion
- Assume that the grandson will die if he treats the body “with respect”.
- 1. Does he have a moral obligation to keep his promise?
- 2. Should he have promised in the first place?
- 3. Should the grandfather have asked him to make that promise, knowing the consequences for his grandson?



Crime and punishment

- Punishment produces utility through:
- Rehabilitation
- Deterrence
- Vengeance

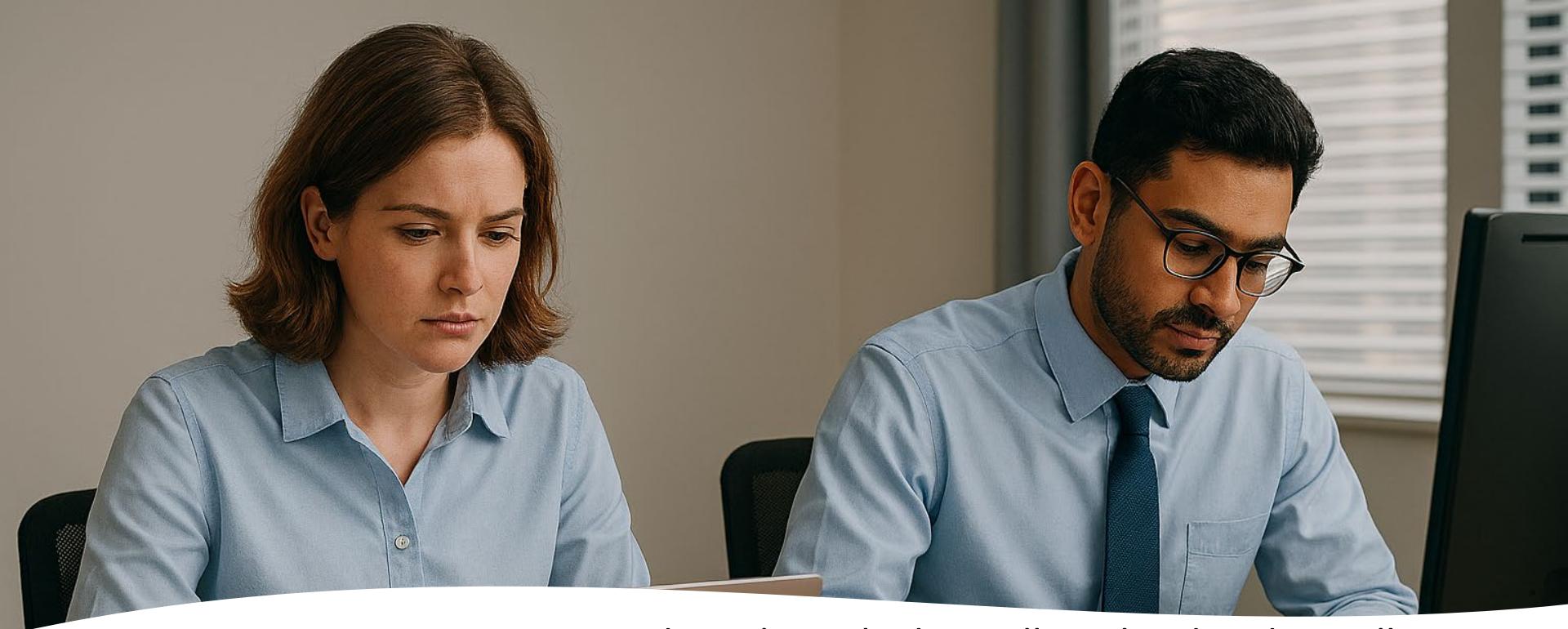


Punishment - problems

- **Rehabilitation**
- Not seen as a right (good in itself) from the point of view of utilitarianism. If we happen to maximize utility by leaving the person e.g. in solitary confinement for life, so be it.
- **Deterrence**
- Why not frame an innocent and punish him in public? Or convict three people, to make sure the one guilty person is punished?
- **Vengeance**
- It can cause pleasure, so any punishment that gives pleasure to enough people is morally justified. Whether the person committed the crime or not is not intrinsically morally relevant for the theory.
- **All these problems derive from not seeing justice and a defense of human rights as values in themselves. Acting on them will have only derivative value when considering utility (e.g. protect someone's freedom of speech if doing that maximizes utility in one particular instance).**

Justice and fairness

- Q: Peter and Sarah worked equally as hard and equally as long. They get paid a single amount to split. Should they split it 50-50?
- A: For utilitarianism, this should be done if it maximizes utility. But if Peter is sad unless he gets 60% of the amount, this would factor in the calculation.
- Because in utilitarianism there is:
- Equal consideration of interests, but...
- Utility is aggregated



Impartiality and Supererogation – Singer’s “Famine, Affluence, and Morality”

- Singer – Must we help Bengali refugees? What obligations do we in affluent nations have toward victims of natural disasters or other disasters of that magnitude?
- Singer’s answer: We must be impartial (emotional proximity and physical distance do not matter to utilitarianism).
- Singer: It’s also **not mere *charity*** to help (charity defined as beyond the call of duty – ethical, but not strictly *required*). It’s our **duty** – definitely required!
- Note that these ideas can easily be applied to computing and are obviously not specific to natural disasters, refugees, etc:
- You’re asked to implement a dark pattern that impacts users in India. How far you go to protect those users?
- Should you point out to your boss that this in fact a dark pattern? Does it depend on how much is at stake for you?

Impartiality

- Singer: “[W]e cannot discriminate against someone merely because he is far away from us” (in distance or psychologically – friend or stranger)
- Discussion
- Three kids fall in the water. You can save two kids who are strangers or your own kid. Must you save the two strangers?
- What if it was 11 strangers versus your family of 10 (partner, grandparents, parents, kids)? Is it your moral obligation to let your whole family die?
- A1. Absolutely. Impartiality is the only way to be moral.
- A2. Possible objection against utilitarianism: “No ought without can”. Almost no one *can be* that impartial. So there’s no obligation to be that impartial.
- A3. Possible objection against utilitarianism: It’s actually wrong to be that impartial. There are special duties to those who rely on us (e.g. our children, partner, friends).

Singer on supererogation – Actions beyond the call of duty

- “It is beyond the scope of my argument to consider whether the distinction [between duty and charity] should be **redrawn** or **abolished** altogether”
- In other words, Singer proposes to **change** or **eliminate** the threshold of supererogation. He seems to propose here the more reasonable **redrawing**:
- “The strong version, which required us to prevent bad things from happening unless in doing so we would be sacrificing something of comparable moral significance, does seem to require reducing ourselves to the level of marginal utility. I should also say that the strong version seems to me to be the correct one. I proposed the more moderate version - that we should prevent bad occurrences unless, to do so, we had to sacrifice something morally significant” Singer

The threshold of supererogation

- If it's **redrawn**, then we need to speak of a moral duty to **do some charity**
- If it's **abolished**, no duty is considered supererogatory (above and beyond the call of duty) – note that this what classical act utilitarianism actually calls for!
- A consequence of abolishing the threshold, which Singer and classical act utilitarianism say is ultimately right – we can no longer distinguish between:
 - **Duty** (it is our duty to give - ethically mandatory)
 - **charity** (usually considered to be within the realm of generosity above the call of duty – moral to give, but optional)

Threshold of supererogation

Do we move it up (moderate proposal) or eliminate it (utilitarianism proper)?

Strong proposal – always maximize utility
(actual utilitarianism)

Charity (moral but optional)

Threshold of supererogation

Duty (moral - must be done)

Discussion (time permitting) Or consider this at home

If you still want to be a utilitarian, you should have responses to the objections raised.

1. Consider the objections we discussed (see a list on the next slide)
2. Do you have responses for each? Identify the most problematic objection, and come up with an example to illustrate how its problematic
3. In light of your responses, ask yourself whether you are a utilitarian?
4. Can we entirely give up on considerations of social utility? (given enough social utility, does it ever tip the scale, where we should be prepared to engage in some actions even if they seem to violate individual rights?)

Popular objections to utilitarianism

- It's a doctrine of swine
- Lack of time to calculate utility
- Too much for humanity
- Motivational/psychological hedonism seems to be mistaken
- Normative hedonism seems to be mistaken
- Desire-fulfillment theory of welfare (Mill's proof) doesn't work
- Aggregating utility often leads to counter-intuitive results
- Strong doctrine of negative responsibility is mistaken
- All sorts of trivial actions are incorrectly classified as morally wrong
- Keeping promises seems of little importance
- Punishment doesn't need to be fair
- Other issues with justice (i.e. fairness) seems of little importance in many situations
- Counter-intuitive view of supererogatory action
- Demand for perfect impartiality counter-intuitive (demands seem inappropriate or even impossible)