# 1NC vs Memorial DX

## OFF

### 1NC - OFF

Con Con CP

#### CP Text: States ought to call a global constitutional convention and establish a constitution reflecting intergenerational concern with exclusive authority to the plan and bind participating bodies to its result.

#### The CP applies intergenerational equity to future generations – that’s better than trying to decide now whether the plan is beneficial across deep time – every country would say yes.

Tan 2k [David Tan, LL.M., Harvard Law School; LL.B. (Hons), B.Com., University of Melbourne. Former Tutor in Law, Trinity College, University of Melbourne, “Towards a New Regime for the Protection of Outer Space as the "Province of All Mankind",” 2000, *The Yale Journal of International Law*, Vol. 25, https://digitalcommons.law.yale.edu/cgi/viewcontent.cgi?article=1114&context=yjil]

Edith Brown Weiss has advanced the theory of “intergenerational equity,” which provides for generational rights and obligations.158 Her thesis consists of a normative framework of intersecting theories of intergenerational and intragenerational equity that are derived from an underlying planetary trust, embodying the notion that generations act as stewards to sustain the welfare and well-being of all generations. This planetary trust obliges “each generation to preserve the diversity of the resource base and to pass the planet to future generations in no worse condition than it receives it.”159 The principle of the conservation of options requires each generation “to conserve the diversity of the natural and cultural resource base, so that it does not unduly restrict the options available to future generations in solving their problems and satisfying their own values, and should be entitled to diversity comparable to that enjoyed by previous generations.”\*60 The theory of intergenerational equity is an appealing one. Unfortunately, Weiss’s model generally rests upon an intertemporal human rights model for preserving the global environment. This presents many problems, ranging from the questionable existence of the right to a decent environment to the issue of remedies in respect of claims made by future generations against present generations.161

Whether the global awareness of the harm to our sense of intergenerational identity, as evidenced by the various U.N. General Assembly resolutions and numerous international conventions, will be sufficient to mobilize the implementation and enforcement of effective legal measures on behalf of future generations is doubtful. But more importantly, the notions of intergenerational identity and sustainable development will prove to be invaluable concepts in framing the discussion in Part VI.

Current literature has concentrated on the notion of sustainable development as involving the integration of economic and environmental considerations at all levels of decision-making.162 But the outer-space environment has been largely ignored, as if it were simply economic development on Earth that must be environmentally sound. There is no reason, however, why the precautionary principles that emerge from the concept of sustainable development in the Stockholm Declaration, the Rio Declaration, and the World Charter for Nature should not apply equally to the outer-space environment. Few states, if any, will take issue with the proposition that the exploration and use of outer space should be sustainable. It is in the common interest of all states, whether spacefaring or otherwise, to subscribe to a regime that allows for the development of space activities in a manner that leaves the space environment in a substantially unimpaired condition for future generations. One might even ultimately find that the uniqueness and vulnerability of the outer-space environment demand that the international community as a whole recognize sustainable development as a “global ethic”163 that transcends terrestrial boundaries, as a peremptory norm that prohibits “policies and practices that support current living standards by depleting the productive base, including natural resources, and that leaves future generations with poorer prospects and greater risks than our own.”164 We should not confine our actions to those we are now able to determine as directly or indirectly benefiting ourselves or our descendants. On the contrary, we should “cultivate our natural sense of obligation not to act wastefully or wantonly even when we cannot calculate how such acts would make any present or future persons worse off.”165 It seems impossible to find universally agreed-upon limits on the freedom of exploration and use of outer space. Rather than focus on indeterminate rules of custom-formation, we should concentrate on establishing fair and workable arrangements and institutions that can successfully accommodate the competing interests of all nations. With these guidelines in mind, we will now examine new methods of treaty-making that will enhance the willingness of states to participate in an environmental program that seeks to achieve an acceptable balance between pollution control and freedom of space exploration.

#### That solves the aff – it addresses shared anxieties while building political consensus.

Gardiner 14 1 [Stephen M. Gardiner, Professor of Philosophy and Ben Rabinowitz Endowed Professor of Human Dimensions of the Environment at the University of Washington, Seattle, “A Call for a Global Constitutional Convention Focused on Future Generations,” 2014, *Ethics & International Affairs*, Vol. 28, Issue 3, pp. 299-315, https://doi.org/10.1017/S0892679414000379, EA]

A Constitutional Convention

In my view, the above line of reasoning leads naturally to a more specific proposal: that we—concerned individuals, interested community groups, national governments, and transnational organizations—should initiate a call for a global constitutional convention focused on future generations. This proposal has two components. The first component is procedural. The proposal takes the form of a “call to action.” It is explicitly an attempt to engage a range of actors, based on a claim that they have or should take on a set of responsibilities, and a view about how to go about discharging those responsibilities. The second component is substantive. The main focus for action is a push for the creation of a constitutional convention at the global level, whose role is to pave the way for an overall constitutional system that appropriately embodies intergenerational concern.

The substantive idea rests on several key ideas. Still, for the purposes of a basic proposal, I suggest that these be understood in a relatively open way that, as far as is practicable, does not prejudge the outcome of the convention, and especially its main recommendations. First, the convention itself should be understood as “a representative body called together for some occasional or temporary purpose” and “constituted by statute to represent the people in their primary relations.”14 Second, a constitutional system should be thought of in a minimalist sense as “a set of norms (rules, principles or values) creating, structuring, and possibly defining the limits of government power or authority.”15 Third, the “instigating” role of the convention should be to discuss, develop, make recommendations toward, and set in motion a process for the establishment of a constitution. Fourth, its primary subject matter should be the need to adequately reflect and embody intergenerational concern, where this would include at least the protection of future generations, the promotion of their interests (where “interests” is to be broadly conceived so as to include rights, claims, welfare, and so on), and the discharging of duties with respect to them. It may also (and in my view should) include some way of reflecting concern for past generations, including responsiveness to at least certain of their interests and views. However, I will leave that issue aside in what follows.

The proposal to initiate a call for a global constitutional convention has at least two attractive features. First, it is based in a deep political reality, and does not underplay the challenge. It acknowledges the problem as it is, both specific and general, and calls attention to the heart of that problem, including to the failures of the current system, the need for an alternative, and the background issue of responsibility. Moreover, though the proposal is dramatic and rhetorically eye-catching, it is so in a way that is appropriately responsive to the seriousness of the issue at hand, the persistent political inertia surrounding more modest initiatives, and the fact that (grave though concerns about it are) climate change is only one instance of the tyranny of the contemporary (and the wider perfect moral storm), and we should expect others to arise over the coming decades and centuries.

The second attractive feature of the proposal is that, though ambitious, it is not alienating. While it does not succumb to despair in the face of the challenge, neither does it needlessly polarize and divide from the outset (for example, by leaping to specific recommendations about how to fill the institutional gap). Instead, it acknowledges that there are fundamental difficulties and anxieties, but uses them to start the right kind of debate, rather than to foreclose it. As a result, the proposal is a promising candidate to serve as the subject of a wide and overlapping political consensus, at least among those who share intergenerational concern.

Selective Mirroring

To quell some initial anxieties, it is perhaps worth clarifying the open-ended and non-alienating character of the proposal. One temptation would be to view the call for a global constitutional convention as a fairly naked plea for world government, a prospect that would be deeply alienating—indeed anathema—to many. However, that is not my intention. Though it is possible that a global constitutional convention would lead in this direction, it is by no means certain.

At a minimum, no such body could plausibly recommend any form of “world government” without simultaneously advancing detailed suggestions about how to avoid the standard threats such an institution might pose. Moreover, it seems perfectly conceivable, even likely under current ways of thinking, that a global constitutional convention would pursue what we might call a selective mirroring strategy. Specifically, a convention would seek to develop a broader system of institutions and practices that reflected the desirable features of a powerful and highly centralized global authority but neutralized the standing threats posed by it (for example, it might employ familiar strategies such as the separation of powers). In all likelihood, one feature of a selective mirroring approach would be the significant preservation of existing institutions to serve as a bulwark against the excesses of any newly created ones. Whether and how such a strategy might be made effective against the perfect moral storm, and whether something closer to a “world government” would do better, would be a central issue for discussion by the convention.

#### It spills over to foster broader intergenerational representation, but independence is key

Gardiner 14 2 [Stephen M. Gardiner, Professor of Philosophy and Ben Rabinowitz Endowed Professor of Human Dimensions of the Environment at the University of Washington, Seattle, “A Call for a Global Constitutional Convention Focused on Future Generations,” 2014, *Ethics & International Affairs*, Vol. 28, Issue 3, pp. 299-315, https://doi.org/10.1017/S0892679414000379, EA]

One set of guidelines concerns how the global constitutional convention relates to other institutions. The first guideline concerns relative independence:

(1) Autonomy: Any global constitutional convention should have considerable autonomy from other institutions, and especially from those dominated by factors that generate or facilitate the tyranny of the contemporary (and the perfect moral storm, more generally).

Thus, for example, attempts should be made to insulate the global constitutional convention from too much influence from short-term and narrowly economic forces.

The second guideline concerns limits to that independence:

(2) Mutual Accountability: Any global constitutional convention should be to some extent accountable to other major institutions, and they should be accountable to it.

Thus, for example, though the global constitutional convention should not be able to decide unilaterally that national institutions should be radically supplanted, nevertheless such institutions should not have a simple veto on the recommendations of the convention, including those that would result in sharp limits to their powers.

A third guideline concerns adequacy:

(3) Functional Adequacy: The global constitutional convention should be constructed in such a way that it is highly likely to produce recommendations that are functionally adequate to the task.

Thus, for example, the tasks of the global constitutional convention should not be assigned to any currently existing body whose design and authority is clearly unsuitable. In my view, this guideline rules out proposals such as the Royal Society’s suggestion that governance of geoengineering should be taken up by the United Nations’ Commission on Sustainable Development,20 or the Secretary-General’s recommendation of a new United Nations’ High Commissioner for Future Generations.21 Though such proposals may have merit for some purposes (for example, as pragmatic, incremental suggestions to highlight the importance of intergenerational issues), they are too modest, in my opinion, to reflect the gravity of the threats posed by climate change in particular, and the perfect moral storm more generally.

Aims

A second set of guidelines concerns the aims of the global constitutional convention. Here, the perfect moral storm analysis would suggest:

(4) Comprehensiveness: The convention should be under a mandate to consider a very broad range of global, intergenerational issues, to focus on such issues at a foundational level, and to recommend institutional reform accordingly.

(5) Standing Authority: Though the convention may recommend the establishment of some temporary and issue-specific bodies, its focus should be on the establishment of institutions with standing authority over the long term.

These guidelines are significant in that they stand against existing issue-specific approaches to global and intergenerational problems, and encourage not only a less ad hoc but also a more proactive approach. In particular, the global constitutional convention might be expected to recommend institutions that would be charged with identifying, monitoring, and taking charge of intergenerational issues as such. For example, such institutions should address not only specific policy issues (such as climate change, large asteroid detection, and long-term nuclear waste) but also the need to identify similar threats before they arise.

#### Proactive measures mitigate a laundry list of emerging catastrophic risks – extinction.

Beckstead 14 [Nick Beckstead, Nick Bostrom, Niel Bowerman, Owen Cotton-Barratt, William MacAskill, Seán Ó hÉigeartaigh, Toby Ord, \* Future of Humanity Institute, University of Oxford, \*\* Director, Future of Humanity Institute, University of Oxford, \*\*\* Global Priorities Project, Centre for Effective Altruism; Department of Physics, University of Oxford, \*\*\*\* Global Priorities Project, Centre for Effective Altruism; Future of Humanity Institute, University of Oxford, \*\*\*\*\* Uehiro Centre for Practical Ethics, University of Oxford, \*\*\*\*\*\* Cambridge Centre for the Study of Existential Risk; Future of Humanity Institute, University of Oxford, \*\*\*\*\*\*\* Programme on the Impacts of Future Technology, Oxford Martin School, University of Oxford, “Policy Brief: Unprecedented Technological Risks,” 2014, *The Global Priorities Project, The Future of Humanity Institute, The Oxford Martin Programme on the Impacts of Future Technology, and The Centre for the Study of Existential Risk*, https://www.fhi.ox.ac.uk/wp-content/uploads/Unprecedented-Technological-Risks.pdf, Accessed: 03/13/21, EA]

In the near future, major technological developments will give rise to new unprecedented risks. In particular, like nuclear technology, developments in synthetic biology, geoengineering, distributed manufacturing and artificial intelligence create risks of catastrophe on a global scale. These new technologies will have very large benefits to humankind. But, without proper regulation, they risk the creation of new weapons of mass destruction, the start of a new arms race, or catastrophe through accidental misuse. Some experts have suggested that these technologies are even more worrying than nuclear weapons, because they are more difficult to control. Whereas nuclear weapons require the rare and controllable resources of uranium-235 or plutonium-239, once these new technologies are developed, they will be very difficult to regulate and easily accessible to small countries or even terrorist groups.

Moreover, these risks are currently underregulated, for a number of reasons. Protection against such risks is a global public good and thus undersupplied by the market. Implementation often requires cooperation among many governments, which adds political complexity. Due to the unprecedented nature of the risks, there is little or no previous experience from which to draw lessons and form policy. And the beneficiaries of preventative policy include people who have no sway over current political processes — our children and grandchildren.

Given the unpredictable nature of technological progress, development of these technologies may be unexpectedly rapid. A political reaction to these technologies only when they are already on the brink of development may therefore be too late. We need to implement prudent and proactive policy measures in the near future, even if no such breakthroughs currently appear imminent.

### 1NC - OFF

Util NC

**Only consequentialism is epistemically accessible ---**

#### Physical facts influence morality

**Papineau** David [Professor of Philosophy King's College London], First published Thu Feb 22, 2007; substantive revision Tue Mar 31, 2020 https://plato.stanford.edu/entries/naturalism/#MorFac

Moore took this argument to show that moral facts constitute a distinct species of non-natural fact. However, any such non-naturalist view of morality faces immediate difficulties, deriving ultimately from the kind of causal closure thesis discussed above. If all physical effects are due to a limited range of physically-grounded natural causes, and if moral facts lie outside this range, then it follow that moral facts can never make any difference to what happens in the physical world (Harman 1986). At first sight this may seem tolerable (perhaps moral facts indeed don’t have any physical effects). But it has awkward epistemological consequences. For beings like us, knowledge of the spatiotemporal world is mediated by physical processes involving our sense organs and cognitive systems. If moral facts cannot influence the physical world, then it is hard to see how we can have any knowledge of them. The traditional non-naturalist answer to this problem is to posit a non-natural faculty of “moral intuition” that gives us some kind of direct access to the moral realm (as explained in Ridge 2014: Section 3). However, causal closure once more makes it difficult to make good sense of this suggestion. Presumably at some point the posited intuitive faculty will need to make a causal difference in the physical world (by affecting what people say and do, for example). And at this point the causal closure argument will bite once more, to show that a non-natural intuitive faculty would implausibly imply that some of our actions are strongly overdetermined by two metaphysically independent antecedents. Moral non-naturalism has had something of a revival in recent years, with defenders including Russ Shaffer-Landau (2003), Ralph Wedgwood (2007), Derek Parfit (2011) and David Enoch (2011). Still, the challenge of accounting for our access to non-natural moral facts remains, and it is debatable whether any of these writers has found a satisfactory alternative to a causally problematic faculty of intuition. Perhaps the most developed suggestion is Enoch’s (2011) appeal to the indispensability of non-natural moral facts to moral reasoning, a line of argument that is analogous to Hilary Putnam’s case for non-natural mathematical objects, to be discussed in the next section below. But Enoch’s appeal arguably faces many of the same general objections as Putnam’s argument, as well as objections specific to the moral realm (see Leng 2016). In light of the difficulties facing moral non-naturalism, most contemporary moral philosophers opt instead for some species of naturalist view. We can divide the naturalist options here into two broad categories: irrealist and realist. Irrealist moral naturalists aim to account for moral discourse by offering naturalist accounts of the social and linguistic and practices that govern it, but without supposing that moral utterances report on moral facts with a substantial independent existence (Joyce 2015). By contrast, naturalist moral realists agree with moral non-naturalists that substantial moral facts exist, but seek to locate them in the natural realm rather than in some sui generis non-natural realm (Lenman 2014). Both these broad categories have further sub-divisions. Among the irrealists, we can distinguish explicitly non-cognitivist views like emotivism and prescriptivism which deny that moral judgements express beliefs (Hare 1952, Blackburn 1993, Gibbard 2003) from cognitivist views that accept that moral judgements do express beliefs but deny a substantial reality to the putative facts to which they answer; and among the latter cognitivist views we can distinguish error-theoretic fictionalist options which view moral judgements as simply false (Mackie 1977, Kalderon 2005) from projectivist options which hold that moral discourse is sufficiently disciplined for its judgements to qualify for a species of truth even though they do not report on independently existing causally significant facts (Wright 1992, Price 2011). Naturalist moral realism also comes in different varieties. In recent debates two versions have figured prominently; “Cornell realism”, which includes moral facts among the causally significant facts but resists their type-reducibility to non-moral facts (Sturgeon 1985, Boyd 1988), and “moral functionalism” which is happy to equate moral facts with straightforwardly descriptive facts (Jackson 1998). Any kind of moral naturalist realist needs to reject Moore’s open question argument. There are two alternatives here. One is to insist that Moore’s posited openness is relatively superficial, and that there is no principled barrier to inferring moral facts a priori from the non-moral natural facts, even if such inferences will sometimes require a significant amount of information and reflection. The other is to argue that the constitution of moral facts by non-moral natural facts is an a posteriori matter, akin to the relation between water and H2O, and that therefore Moore’s openness only points to a conceptual gap, not a metaphysical one (Ridge 2014: Section 2).

#### The normative supervenes on the natural – natural facts like whether brains develop to permit rationality or subjectivity determine whether non naturalist moral facts can be premised on things like capacity for reason

**Lutz and Lenman 18.** Lutz, Matthew and Lenman, James, "Moral Naturalism", The Stanford Encyclopedia of Philosophy (Fall 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2018/entries/naturalism-moral/>. //Massa

The first argument against normative non-naturalism concerns normative supervenience. **The normative supervenes on the natural; in all** metaphysically **possible worlds in which the natural facts are the same as** they are in **the actual world, the moral facts are the same** as well. **This** claim **has been called the “least controversial thesis in metaethics”** (Rosen forthcoming); **it is very widely accepted.** But it is also a striking fact that stands in need of some explanation. **For naturalists**, such an explanation is easy to provide: **the moral facts just are natural facts, so when we consider worlds that are naturally the same** as the actual world, **we will ipso facto be considering worlds that are morally the same** as the actual world. But for the non-naturalist, no such explanation seems available. In fact, **it seems** to be in principle **impossible for a non-naturalist to explain how the moral supervenes on the natural.** And if the non-naturalist can offer no explanation of this phenomenon that demands explanation, this is a heavy mark against non-naturalism (McPherson 2012).

#### Pleasure and pain are intrinsic value and disvalue --- two warrants:

1. **Everything else *regresses* – robust neuroscience.**

**Blum et al. 18**

Kenneth Blum, 1Department of Psychiatry, Boonshoft School of Medicine, Dayton VA Medical Center, Wright State University, Dayton, OH, USA 2Department of Psychiatry, McKnight Brain Institute, University of Florida College of Medicine, Gainesville, FL, USA 3Department of Psychiatry and Behavioral Sciences, Keck Medicine University of Southern California, Los Angeles, CA, USA 4Division of Applied Clinical Research & Education, Dominion Diagnostics, LLC, North Kingstown, RI, USA 5Department of Precision Medicine, Geneus Health LLC, San Antonio, TX, USA 6Department of Addiction Research & Therapy, Nupathways Inc., Innsbrook, MO, USA 7Department of Clinical Neurology, Path Foundation, New York, NY, USA 8Division of Neuroscience-Based Addiction Therapy, The Shores Treatment & Recovery Center, Port Saint Lucie, FL, USA 9Institute of Psychology, Eötvös Loránd University, Budapest, Hungary 10Division of Addiction Research, Dominion Diagnostics, LLC. North Kingston, RI, USA 11Victory Nutrition International, Lederach, PA., USA 12National Human Genome Center at Howard University, Washington, DC., USA, Marjorie Gondré-Lewis, 12National Human Genome Center at Howard University, Washington, DC., USA 13Departments of Anatomy and Psychiatry, Howard University College of Medicine, Washington, DC US, Bruce Steinberg, 4Division of Applied Clinical Research & Education, Dominion Diagnostics, LLC, North Kingstown, RI, USA, Igor Elman, 15Department Psychiatry, Cooper University School of Medicine, Camden, NJ, USA, David Baron, 3Department of Psychiatry and Behavioral Sciences, Keck Medicine University of Southern California, Los Angeles, CA, USA, Edward J Modestino, 14Department of Psychology, Curry College, Milton, MA, USA, Rajendra D Badgaiyan, 15Department Psychiatry, Cooper University School of Medicine, Camden, NJ, USA, Mark S Gold 16Department of Psychiatry, Washington University, St. Louis, MO, USA, “Our evolved unique pleasure circuit makes humans different from apes: Reconsideration of data derived from animal studies”, U.S. Department of Veterans Affairs, 28 February 2018, accessed: 19 August 2020, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6446569/>, R.S.

**Pleasure** is not only one of the three primary reward functions but it also **defines reward.** As homeostasis explains the functions of only a limited number of rewards, the principal reason why particular stimuli, objects, events, situations, and activities are rewarding may be due to pleasure. This applies first of all to sex and to the primary homeostatic rewards of food and liquid and extends to money, taste, beauty, social encounters and nonmaterial, internally set, and intrinsic rewards. Pleasure, as the primary effect of rewards, drives the prime reward functions of learning, approach behavior, and decision making and provides the **basis for hedonic theories** of reward function. We are attracted by most rewards and exert intense efforts to obtain them, just because they are enjoyable [10].

Pleasure is a passive reaction that derives from the experience or prediction of reward and may lead to a long-lasting state of happiness. The word happiness is difficult to define. In fact, just obtaining physical pleasure may not be enough. One key to happiness involves a network of good friends. However, it is not obvious how the higher forms of satisfaction and pleasure are related to an ice cream cone, or to your team winning a sporting event. Recent multidisciplinary research, using both humans and detailed invasive brain analysis of animals has discovered some critical ways that the brain processes pleasure [14].

Pleasure as a hallmark of reward is sufficient for defining a reward, but it may not be necessary. A reward may generate positive learning and approach behavior simply because it contains substances that are essential for body function. When we are hungry, we may eat bad and unpleasant meals. A monkey who receives hundreds of small drops of water every morning in the laboratory is unlikely to feel a rush of pleasure every time it gets the 0.1 ml. Nevertheless, with these precautions in mind, we may define any stimulus, object, event, activity, or situation that has the potential to produce pleasure as a reward. In the context of reward deficiency or for disorders of addiction, homeostasis pursues pharmacological treatments: drugs to treat drug addiction, obesity, and other compulsive behaviors. The theory of allostasis suggests broader approaches - such as re-expanding the range of possible pleasures and providing opportunities to expend effort in their pursuit. [15]. It is noteworthy, the first animal studies eliciting approach behavior by electrical brain stimulation interpreted their findings as a discovery of the brain’s pleasure centers [16] which were later partly associated with midbrain dopamine neurons [17–19] despite the notorious difficulties of identifying emotions in animals.

Evolutionary theories of pleasure: The love connection BO:D

Charles Darwin and other biological scientists that have examined the biological evolution and its basic principles found various mechanisms that steer behavior and biological development. Besides their theory on natural selection, it was particularly the sexual selection process that gained significance in the latter context over the last century, especially when it comes to the question of what makes us “what we are,” i.e., human. However, the capacity to sexually select and evolve is not at all a human accomplishment alone or a sign of our uniqueness; yet, we humans, as it seems, are ingenious in fooling ourselves and others–when we are in love or desperately search for it.

It is well established that modern biological theory conjectures that **organisms are** the **result of evolutionary competition.** In fact, Richard Dawkins stresses gene survival and propagation as the basic mechanism of life [20]. Only genes that lead to the fittest phenotype will make it. It is noteworthy that the phenotype is selected based on behavior that maximizes gene propagation. To do so, the phenotype must survive and generate offspring, and be better at it than its competitors. Thus, the ultimate, distal function of rewards is to increase evolutionary fitness by ensuring the survival of the organism and reproduction. It is agreed that learning, approach, economic decisions, and positive emotions are the proximal functions through which phenotypes obtain other necessary nutrients for survival, mating, and care for offspring.

Behavioral reward functions have evolved to help individuals to survive and propagate their genes. Apparently, people need to live well and long enough to reproduce. Most would agree that homo-sapiens do so by ingesting the substances that make their bodies function properly. For this reason, foods and drinks are rewards. Additional rewards, including those used for economic exchanges, ensure sufficient palatable food and drink supply. Mating and gene propagation is supported by powerful sexual attraction. Additional properties, like body form, augment the chance to mate and nourish and defend offspring and are therefore also rewards. Care for offspring until they can reproduce themselves helps gene propagation and is rewarding; otherwise, many believe mating is useless. According to David E Comings, as any small edge will ultimately result in evolutionary advantage [21], additional reward mechanisms like novelty seeking and exploration widen the spectrum of available rewards and thus enhance the chance for survival, reproduction, and ultimate gene propagation. These functions may help us to obtain the benefits of distant rewards that are determined by our own interests and not immediately available in the environment. Thus the distal reward function in gene propagation and evolutionary fitness defines the proximal reward functions that we see in everyday behavior. That is why foods, drinks, mates, and offspring are rewarding.

There have been theories linking pleasure as a required component of health benefits salutogenesis, (salugenesis). In essence, under these terms, pleasure is described as a state or feeling of happiness and satisfaction resulting from an experience that one enjoys. Regarding pleasure, it is a double-edged sword, on the one hand, it promotes positive feelings (like mindfulness) and even better cognition, possibly through the release of dopamine [22]. But on the other hand, pleasure simultaneously encourages addiction and other negative behaviors, i.e., motivational toxicity. It is a complex neurobiological phenomenon, relying on reward circuitry or limbic activity. It is important to realize that through the “Brain Reward Cascade” (BRC) endorphin and endogenous morphinergic mechanisms may play a role [23]. While natural rewards are essential for survival and appetitive motivation leading to beneficial biological behaviors like eating, sex, and reproduction, crucial social interactions seem to further facilitate the positive effects exerted by pleasurable experiences. Indeed, experimentation with addictive drugs is capable of directly acting on reward pathways and causing deterioration of these systems promoting hypodopaminergia [24]. Most would agree that pleasurable activities can stimulate personal growth and may help to induce healthy behavioral changes, including stress management [25]. The work of Esch and Stefano [26] concerning the link between compassion and love implicate the brain reward system, and pleasure induction suggests that social contact in general, i.e., love, attachment, and compassion, can be highly effective in stress reduction, survival, and overall health.

Understanding the role of neurotransmission and pleasurable states both positive and negative have been adequately studied over many decades [26–37], but comparative anatomical and neurobiological function between animals and homo sapiens appear to be required and seem to be in an infancy stage.

Finding happiness is different between apes and humans

As stated earlier in this expert opinion one key to happiness involves a network of good friends [38]. However, it is not entirely clear exactly how the higher forms of satisfaction and pleasure are related to a sugar rush, winning a sports event or even sky diving, all of which augment dopamine release at the reward brain site. Recent multidisciplinary research, using both humans and detailed invasive brain analysis of animals has discovered some critical ways that the brain processes pleasure.

Remarkably, there are pathways for ordinary liking and pleasure, which are limited in scope as described above in this commentary. However, there are **many brain regions**, often termed hot and cold spots, that significantly **modulate** (increase or decrease) our **pleasure or** even produce **the opposite** of pleasure— that is disgust and fear [39]. One specific region of the nucleus accumbens is organized like a computer keyboard, with particular stimulus triggers in rows— producing an increase and decrease of pleasure and disgust. Moreover, the cortex has unique roles in the cognitive evaluation of our feelings of pleasure [40]. Importantly, the interplay of these multiple triggers and the higher brain centers in the prefrontal cortex are very intricate and are just being uncovered.

Desire and reward centers

It is surprising that many different sources of pleasure activate the same circuits between the mesocorticolimbic regions (Figure 1). Reward and desire are two aspects pleasure induction and have a very widespread, large circuit. Some part of this circuit distinguishes between desire and dread. The so-called pleasure circuitry called “REWARD” involves a well-known dopamine pathway in the mesolimbic system that can influence both pleasure and motivation.

In simplest terms, the well-established mesolimbic system is a dopamine circuit for reward. It starts in the ventral tegmental area (VTA) of the midbrain and travels to the nucleus accumbens (Figure 2). It is the cornerstone target to all addictions. The VTA is encompassed with neurons using glutamate, GABA, and dopamine. The nucleus accumbens (NAc) is located within the ventral striatum and is divided into two sub-regions—the motor and limbic regions associated with its core and shell, respectively. The NAc has spiny neurons that receive dopamine from the VTA and glutamate (a dopamine driver) from the hippocampus, amygdala and medial prefrontal cortex. Subsequently, the NAc projects GABA signals to an area termed the ventral pallidum (VP). The region is a relay station in the limbic loop of the basal ganglia, critical for motivation, behavior, emotions and the “Feel Good” response. This defined system of the brain is involved in all addictions –substance, and non –substance related. In 1995, our laboratory coined the term “Reward Deficiency Syndrome” (RDS) to describe genetic and epigenetic induced hypodopaminergia in the “Brain Reward Cascade” that contribute to addiction and compulsive behaviors [3,6,41].

Furthermore, ordinary “liking” of something, or pure pleasure, is represented by small regions mainly in the limbic system (old reptilian part of the brain). These may be part of larger neural circuits. In Latin, hedus is the term for “sweet”; and in Greek, hodone is the term for “pleasure.” Thus, the word Hedonic is now referring to various subcomponents of pleasure: some associated with purely sensory and others with more complex emotions involving morals, aesthetics, and social interactions. The capacity to have pleasure is part of being healthy and may even extend life, especially if linked to optimism as a dopaminergic response [42].

Psychiatric illness often includes symptoms of an abnormal inability to experience pleasure, referred to as anhedonia. A negative feeling state is called dysphoria, which can consist of many emotions such as pain, depression, anxiety, fear, and disgust. Previously many scientists used animal research to uncover the complex mechanisms of pleasure, liking, motivation and even emotions like panic and fear, as discussed above [43]. However, as a significant amount of related research about the specific brain regions of pleasure/reward circuitry has been derived from invasive studies of animals, these cannot be directly compared with subjective states experienced by humans.

In an attempt to resolve the controversy regarding the causal contributions of mesolimbic dopamine systems to reward, we have previously evaluated the three-main competing explanatory categories: “liking,” “learning,” and “wanting” [3]. That is, dopamine may mediate (a) liking: the hedonic impact of reward, (b) learning: learned predictions about rewarding effects, or (c) wanting: the pursuit of rewards by attributing incentive salience to reward-related stimuli [44]. We have evaluated these hypotheses, especially as they relate to the RDS, and we find that the incentive salience or “wanting” hypothesis of dopaminergic functioning is supported by a majority of the scientific evidence. Various neuroimaging studies have shown that anticipated behaviors such as sex and gaming, delicious foods and drugs of abuse all affect brain regions associated with reward networks, and may not be unidirectional. Drugs of abuse enhance dopamine signaling which sensitizes mesolimbic brain mechanisms that apparently evolved explicitly to attribute incentive salience to various rewards [45].

Addictive substances are voluntarily self-administered, and they enhance (directly or indirectly) dopaminergic synaptic function in the NAc. This activation of the brain reward networks (producing the ecstatic “high” that users seek). Although these circuits were initially thought to encode a set point of hedonic tone, it is now being considered to be far more complicated in function, also encoding attention, reward expectancy, disconfirmation of reward expectancy, and incentive motivation [46]. The argument about addiction as a disease may be confused with a predisposition to substance and nonsubstance rewards relative to the extreme effect of drugs of abuse on brain neurochemistry. The former sets up an individual to be at high risk through both genetic polymorphisms in reward genes as well as harmful epigenetic insult. Some Psychologists, even with all the data, still infer that addiction is not a disease [47]. Elevated stress levels, together with polymorphisms (genetic variations) of various dopaminergic genes and the genes related to other neurotransmitters (and their genetic variants), and may have an additive effect on vulnerability to various addictions [48]. In this regard, Vanyukov, et al. [48] suggested based on review that whereas the gateway hypothesis does not specify mechanistic connections between “stages,” and does not extend to the risks for addictions the concept of common liability to addictions may be more parsimonious. The latter theory is grounded in genetic theory and supported by data identifying common sources of variation in the risk for specific addictions (e.g., RDS). This commonality has identifiable neurobiological substrate and plausible evolutionary explanations.

Over many years the controversy of dopamine involvement in especially “pleasure” has led to confusion concerning separating motivation from actual pleasure (wanting versus liking) [49]. We take the position that animal studies cannot provide real clinical information as described by self-reports in humans. As mentioned earlier and in the abstract, on November 23rd, 2017, evidence for our concerns was discovered [50]

In essence, although nonhuman primate brains are similar to our own, the disparity between other primates and those of human cognitive abilities tells us that surface similarity is not the whole story. Sousa et al. [50] small case found various differentially expressed genes, to associate with pleasure related systems. Furthermore, the dopaminergic interneurons located in the human neocortex were absent from the neocortex of nonhuman African apes. Such differences in neuronal transcriptional programs may underlie a variety of neurodevelopmental disorders.

In simpler terms, the system controls the production of dopamine, a chemical messenger that plays a significant role in pleasure and rewards. The senior author, Dr. Nenad Sestan from Yale, stated: “Humans have evolved a dopamine system that is different than the one in chimpanzees.” This may explain why the behavior of humans is so unique from that of non-human primates, even though our brains are so surprisingly similar, Sestan said: “It might also shed light on why people are vulnerable to mental disorders such as autism (possibly even addiction).” Remarkably, this research finding emerged from an extensive, multicenter collaboration to compare the brains across several species. These researchers examined 247 specimens of neural tissue from six humans, five chimpanzees, and five macaque monkeys. Moreover, these investigators analyzed which genes were turned on or off in 16 regions of the brain. While the differences among species were subtle, **there was** a **remarkable contrast in** the **neocortices**, specifically in an area of the brain that is much more developed in humans than in chimpanzees. In fact, these researchers found that a gene called tyrosine hydroxylase (TH) for the enzyme, responsible for the production of dopamine, was expressed in the neocortex of humans, but not chimpanzees. As discussed earlier, dopamine is best known for its essential role within the brain’s reward system; the very system that responds to everything from sex, to gambling, to food, and to addictive drugs. However, dopamine also assists in regulating emotional responses, memory, and movement. Notably, abnormal dopamine levels have been linked to disorders including Parkinson’s, schizophrenia and spectrum disorders such as autism and addiction or RDS.

Nora Volkow, the director of NIDA, pointed out that one alluring possibility is that the neurotransmitter dopamine plays a substantial role in humans’ ability to pursue various rewards that are perhaps months or even years away in the future. This same idea has been suggested by Dr. Robert Sapolsky, a professor of biology and neurology at Stanford University. Dr. Sapolsky cited evidence that dopamine levels rise dramatically in humans when we anticipate potential rewards that are uncertain and even far off in our futures, such as retirement or even the possible alterlife. This may explain what often motivates people to work for things that have no apparent short-term benefit [51]. In similar work, Volkow and Bale [52] proposed a model in which dopamine can favor NOW processes through phasic signaling in reward circuits or LATER processes through tonic signaling in control circuits. Specifically, they suggest that through its modulation of the orbitofrontal cortex, which processes salience attribution, dopamine also enables shilting from NOW to LATER, while its modulation of the insula, which processes interoceptive information, influences the probability of selecting NOW versus LATER actions based on an individual’s physiological state. This hypothesis further supports the concept that disruptions along these circuits contribute to diverse pathologies, including obesity and addiction or RDS.

#### Phenomenal introspection --- one cannot understand what pain without associating it with objective disvalue.

Mendola 06 [Joseph Mendola, (Joseph Mendola is professor and chair in the Department of Philosophy at the University of Nebraska–Lincoln. He is the author of Human Thought and of articles on ethics, metaphysics, and philosophy of mind.) "Goodness and Justice: A Consequentialist Moral Theory" Cambridge University Press, 2006, https://www.cambridge.org/core/books/goodness-and-justice/AE25780DC33533E8797FB684C5FBD36E, DOA:6-7-2019 // WWBW]

While this view is of course controversial in our historical situation, in which many hold that sensory experience is as of yellow though there is nothing in the world that is so, not even a sense datum, or at the very least that the yellow we experience is a natural property constituted by physical properties like a certain range of surface spectral reflectance, still the view in question is, as I’ve said, one live competitor. Indeed, it is often motivated by arguments that are structurally similar to the open-question argument: You look at a gold bar and have a certain sort of phenomenal experience. But it seems to some that it might well be an open question whether your physical twin in a physically identical environment has the same phenomenal experience, or any at all. He might be a zombie or a qualia invert. And the openness of that question suggests to some that **the physical cannot constitute** your **phenomenal experience.** At least such qualia dualism is relatively concrete and robust. Even though it involves physically unconstituted qualia, it involves nothing that is non-natural in Moore’s sense. It is at least concretely comprehensible. And that gives it a great advantage over alternative forms of normative realism. That is my main point, that this so far familiar qualia dualism unexpectedly but very plausibly implies a form of normative realism. **Painfulness** – or, more accurately, the phenomenal property present in certain sorts of extreme and paradigmatic physical pain – **is** a kind of **disvalue**. That is my new idea.34 The phenomenal difference between those in bliss and those in agony includes a difference in a sort of felt phenomenal value. **The phenomenal difference between pain and pleasure seems** (at least in part and sometimes) **to be that the phenomenal component of the former is nastier, intrinsically worse than that of the second. The red knight was stabbed to death.** Just as no one can adequately describe what it was like to be him without capturing his sensation of his red and flowing blood and hence the property of phenomenal redness, so no one can describe what it was like to be him without capturing the nasty sensations he felt and hence the property of phenomenal nastiness or disvalue. And **no one can understand what his phenomenal state was without knowing that it was intrinsically bad, worse than pleasure. No one, not even a Martian, can give a complete and adequate characterization of the red knight’s murder while ignoring the phenomenal state that was a part of that situation. And no one, not even a Martian, can give a complete and adequate characterization of that phenomenal state without capturing its nastiness, its intrinsic disvalue.** The red knight’s murder possessed what we might call objective intrinsic disvalue. If someone feels bad, then there is something bad, at least in cases of extreme physical pain. My further claim, to which constitutive naturalists dissent, is that this involves unconstituted but natural disvalue. **Like other phenomenal properties, the disvalue present in agony is unconstituted by physical properties, though it is itself concrete and natural. It is just like phenomenal yellow.** The objective but unconstituted phenomenal component of agony involves a correspondingly objective and unconstituted phenomenal property that is usually present in cases of at least extreme physical pain, a painfulness or “unpleasant hedonic tone”, as it was once called.35 And **such objective phenomenal properties are, at least in part, a sort of intrinsic disvalue or badness.** Something analogous is true of certain paradigmatic physical pleasures. They involve objective intrinsic value.

#### Thus, the standard is maximizing expected well-being --- specifically hedonistic act util. Prefer additionally ---

#### Adopt a Parliamentary model to account for moral uncertainty. This entails minimizing existential risk.

Bostrom 09 [Bostrom, Nick (*Existential*ist of a different sort). “Moral uncertainty – toward a solution?” 1 January 2009. <http://www.overcomingbias.com/2009/01/moral-uncertainty-towards-a-solution.html>]

It seems people are overconfident about their moral beliefs. But **how should one** reason and **act if one** acknowledges that one **is uncertain about morality** – not just applied ethics but fundamental moral issues? if you don’t know which moral theory is correct? It doesn’t seem **you can[’t] simply plug your uncertainty into expected utility** decision theory and crank the wheel; **because many** moral **theories** state that you **should not** always **maximize** expected **utility.** Even if we limit consideration to consequentialist theories, it still is hard to see how to combine them in the standard decision theoretic framework. For example, suppose you give X% probability to total utilitarianism and (100-X)% to average utilitarianism. Now an action might add 5 utils to total happiness and decrease average happiness by 2 utils. (This could happen, e.g. if you create a new happy person that is less happy than the people who already existed.) Now what do you do, for different values of X? The problem gets even more complicated if we consider not only consequentialist theories but also deontological theories, contractarian theories, virtue ethics, etc. We might even throw various meta-ethical theories into the stew: error theory, relativism, etc. I’m working on a paper on this together with my colleague Toby Ord. We have some arguments against a few possible “solutions” that we think don’t work. On the positive side we have some tricks that work for a few special cases. But beyond that, the best **we have managed** so far is **a** kind of **metaphor, which** we don’t think is literally and exactly correct, and it is a bit under-determined, but it **seems to get things roughly right** and it might point in the right direction: **The Parliamentary Model.** Suppose that you have a set of mutually exclusive moral theories, and that you assign each of these some probability. Now imagine that **each** of these **theorie**s **gets to send** some number of **delegates to The Parliament**. The number of delegates each theory gets to send is **proportional to the probability of the theory.** Then the delegates bargain with one another for support on various issues; and the Parliament reaches a decision by the delegates voting. What you should do is act according to the decisions of this imaginary Parliament. (Actually, we use an extra trick here: we imagine that the delegates act as if the Parliament’s decision were a stochastic variable such that the probability of the Parliament taking action A is proportional to the fraction of votes for A. This has the effect of eliminating the artificial 50% threshold that otherwise gives a majority bloc absolute power. Yet – unbeknownst to the delegates – the Parliament always takes whatever action got the most votes: this way we avoid paying the cost of the randomization!) The idea here is that moral theories get more influence the more probable they are; yet **even a** relatively **weak theory can still get its way on some issues** that the theory think are extremely important **by sacrificing** its influence **on other** i**s**sues that other theories deem more important. For example, **suppose you assign 10% probability to** total **util**itarianism and 90% to moral egoism (just to illustrate the principle). Then **the Parliament** would mostly take actions that maximize egoistic satisfaction; however it **would make some concessions to util**itarianism **on** issues that utilitarianism thinks is especially important. In this example, the person might donate some portion of their income to **existential risks** research and otherwise live completely selfishly. I think there might be wisdom in **this model**. It **avoids the** dangerous and **unstable extremism** that would result **from letting one’s current favorite moral theory completely dictate action**, while still allowing the aggressive pursuit of some non-commonsensical high-leverage strategies so long as they don’t infringe too much on what other major moral theories deem centrally important.

#### Revisionary intuitionism is true and proves util

Yudkowsky 08 [Eliezer Yudkowsky (research fellow of the Machine Intelligence Research Institute; he also writes Harry Potter fan fiction). “The ‘Intuitions’ Behind ‘Utilitarianism.’” 28 January 2008. LessWrong. http://lesswrong.com/lw/n9/the\_intuitions\_behind\_utilitarianism/]

I haven’t said much about metaethics – the nature of morality – because that has a forward dependency on a discussion of the Mind Projection Fallacy that I haven’t gotten to yet. I used to be very confused about metaethics. After my confusion finally cleared up, I did a postmortem on my previous thoughts. I found that my object-level moral reasoning had been valuable and my **meta-level moral reasoning had been** worse than **useless**. And this appears to be a general syndrome – **people do much better when discussing whether torture is** good or **bad than**when they discuss **the meaning of “good” and “bad”. Thus, I deem it prudent to keep moral discussions on the object level** wherever I possibly can. Occasionally people object to any discussion of morality on the grounds that morality doesn’t exist, and in lieu of jumping over the forward dependency to explain that “exist” is not the right term to use here, I generally say, “But what do you do anyway?” and take the discussion back down to the object level. Paul Gowder, though, has pointed out that both the idea of choosing a googolplex dust specks in a googolplex eyes over 50 years of torture for one person, and the idea of “utilitarianism”, depend on “intuition”. He says I’ve argued that the two are not compatible, but charges me with failing to argue for the utilitarian intuitions that I appeal to. Now “intuition” is not how I would describe the computations that underlie human morality and distinguish us, as moralists, from an ideal philosopher of perfect emptiness and/or a rock. But I am okay with using the word “intuition” as a term of art, bearing in mind that “intuition” in this sense is not to be contrasted to reason, but is, rather, the cognitive building block out of which both long verbal arguments and fast perceptual arguments are constructed. **I see** the project of **morality as a project of renormalizing intuition.** We have intuitions about things that seem desirable or undesirable, intuitions about actions that are right or wrong, intuitions about how to resolve conflicting intuitions, intuitions about how to systematize specific intuitions into general principles. **Delete all** the **intuitions, and** you aren’t left with an ideal philosopher of perfect emptiness, **you’re left with a rock. Keep all your** specific **intuitions and** refuse to build upon the reflective ones, and you aren’t left with an ideal philosopher of perfect spontaneity and genuineness, **you’re left with a** grunting **caveperson** running in circles, due to cyclical preferences and similar inconsistencies. “Intuition”, as a term of art, is not a curse word when it comes to morality – there is nothing else to argue from. **Even modus ponens is an “intuition”** in this sense – **it**‘s **just** that modus ponens **still seems like a good idea after being** formalized, **reflected on**, extrapolated out to see if it has sensible consequences, etcetera. So that is “intuition”. However, Gowder did not say what he meant by “utilitarianism”. Does utilitarianism say… That right actions are strictly determined by good consequences? That praiseworthy actions depend on justifiable expectations of good consequences? That probabilities of consequences should normatively be discounted by their probability, so that a 50% probability of something bad should weigh exactly half as much in our tradeoffs? That virtuous actions always correspond to maximizing expected utility under some utility function? That two harmful events are worse than one? That two independent occurrences of a harm (not to the same person, not interacting with each other) are exactly twice as bad as one? That for any two harms A and B, with A much worse than B, there exists some tiny probability such that gambling on this probability of A is preferable to a certainty of B? If you say that I advocate something, or that my argument depends on something, and that it is wrong, do please specify what this thingy is… anyway, I accept 3, 5, 6, and 7, but not 4; I am not sure about the phrasing of 1; and 2 is true, I guess, but phrased in a rather solipsistic and selfish fashion: you should not worry about being praiseworthy. Now, what are the “intuitions” upon which my “utilitarianism” depends? This is a deepish sort of topic, but I’ll take a quick stab at it. First of all, it’s not just that someone presented me with a list of statements like those above, and I decided which ones sounded “intuitive”. Among other things, **if you try to violate** “**util**itarianism”, **you run into paradoxes, contradictions**, circular preferences, **and other** things that aren’t **symptoms of** moral wrongness so much as **moral incoherence**. After you think about moral problems for a while, and also find new truths about the world, and even discover disturbing facts about how you yourself work, you often end up with different moral opinions than when you started out. This does not quite define moral progress, but it is how we experience moral progress. As part of my experienced moral progress, I’ve drawn a conceptual separation between questions of type Where should we go? and questions of type How should we get there? (Could that be what Gowder means by saying I’m “utilitarian”?) The question of where a road goes – where it leads – you can answer by traveling the road and finding out. If you have a false belief about where the road leads, this falsity can be destroyed by the truth in a very direct and straightforward manner. When it comes to wanting to go to a particular place, this want is not entirely immune from the destructive powers of truth. You could go there and find that you regret it afterward (which does not define moral error, but is how we experience moral error). But, even so, wanting to be in a particular place seems worth distinguishing from wanting to take a particular road to a particular place. Our intuitions about where to go are arguable enough, but our intuitions about how to get there are frankly messed up. **After** the two hundred and eighty-seventh **research** study **showing that people will chop their own feet off if you frame the problem the wrong way, you start to distrust first impressions. When you’ve read** enough **research on scope insensitivity** – people will pay only 28% more to protect all 57 wilderness areas in Ontario than one area, **people will pay the same amount to save 50,000 lives as 5,000** lives… that sort of thing… Well, the worst case of scope insensitivity I’ve ever heard of was described here by Slovic: Other recent research shows similar results. Two Israeli psychologists asked people to contribute to a costly life-saving treatment. They could offer that contribution to a group of eight sick children, or to an individual child selected from the group. The target amount needed to save the child (or children) was the same in both cases. Contributions to individual group members far outweighed the contributions to the entire group. There’s other research along similar lines, but I’m just presenting one example, ’cause, y’know, eight examples would probably have less impact. If you know the general experimental paradigm, then the reason for the above behavior is pretty obvious – focusing your attention on a single child creates more emotional arousal than trying to distribute attention around eight children simultaneously. So people are willing to pay more to help one child than to help eight. Now, **you could** look at this intuition, and **think it was** revealing **some** kind of incredibly **deep moral truth** which shows that one child’s good fortune is somehow devalued by the other children’s good fortune. But what about the billions of other children in the world? Why isn’t it a bad idea to help this one child, when that causes the value of all the other children to go down? How can it be significantly better to have 1,329,342,410 happy children than 1,329,342,409, but then somewhat worse to have seven more at 1,329,342,417? **Or you could** look at that and **say: “The intuition is wrong: the brain can’t** successfully **multiply** by eight and get a larger quantity than it started with. **But it ought to**, normatively speaking.” And once you realize that the brain can’t multiply by eight, then the other cases of scope neglect stop seeming to reveal some fundamental truth about 50,000 lives being worth just the same effort as 5,000 lives, or whatever. You don’t get the impression you’re looking at the revelation of a deep moral truth about nonagglomerative utilities. It’s just that the brain doesn’t goddamn multiply. Quantities get thrown out the window. If you have $100 to spend, and you spend $20 each on each of 5 efforts to save 5,000 lives, you will do worse than if you spend $100 on a single effort to save 50,000 lives. Likewise if such choices are made by 10 different people, rather than the same person. As soon as you start believing that it is better to save 50,000 lives than 25,000 lives, that simple preference of final destinations has implications for the choice of paths, when you consider five different events that save 5,000 lives. (It is a general principle that Bayesians see no difference between the long-run answer and the short-run answer; you never get two different answers from computing the same question two different ways. But the long run is a helpful intuition pump, so I am talking about it anyway.) The aggregative valuation strategy of “shut up and multiply” arises from the simple preference to have more of something – to save as many lives as possible – when you have to describe general principles for choosing more than once, acting more than once, planning at more than one time. Aggregation also arises from claiming that the local choice to save one life doesn’t depend on how many lives already exist, far away on the other side of the planet, or far away on the other side of the universe. Three lives are one and one and one. No matter how many billions are doing better, or doing worse. 3 = 1 + 1 + 1, no matter what other quantities you add to both sides of the equation. And if you add another life you get 4 = 1 + 1 + 1 + 1. That’s aggregation. **When you’ve read** enough heuristics and **biases research, and**enough **coherence** and uniqueness **proofs for** Bayesian probabilities and **expected utility**, and you’ve seen the “Dutch book” and “money pump” effects that penalize trying to handle uncertain outcomes any other way, then **you don’t see** the **preference reversals** in the Allais Paradox **as** revealing some incredibly **deep moral truth** about the intrinsic value of certainty. **It** just **goes to show that the brain doesn’t** goddamn **multiply.** The primitive, perceptual intuitions that make a choice “feel good” don’t handle probabilistic pathways through time very skillfully, especially when the probabilities have been expressed symbolically rather than experienced as a frequency. So you reflect, devise more trustworthy logics, and think it through in words. When you see people insisting that no amount of money whatsoever is worth a single human life, and then driving an extra mile to save $10; or when you see people insisting that no amount of money is worth a decrement of health, and then choosing the cheapest health insurance available; then you don’t think that their protestations reveal some deep truth about incommensurable utilities. Part of it, clearly, is that **primitive intuitions don’t**successfully **diminish the emotional impact of** symbols standing for **small quantities** – anything you talk about seems like “an amount worth considering”. And part of it has to do with preferring unconditional social rules to conditional social rules. Conditional rules seem weaker, seem more subject to manipulation. If there’s any loophole that lets the government legally commit torture, then the government will drive a truck through that loophole. So it seems like there should be an unconditional social injunction against preferring money to life, and no “but” following it. Not even “but a thousand dollars isn’t worth a 0.0000000001% probability of saving a life”. Though the latter choice, of course, is revealed every time we sneeze without calling a doctor. The rhetoric of sacredness gets bonus points for seeming to express an unlimited commitment, an unconditional refusal that signals trustworthiness and refusal to compromise. So you conclude that moral rhetoric espouses qualitative distinctions, because espousing a quantitative tradeoff would sound like you were plotting to defect. On such occasions, people vigorously want to throw quantities out the window, and they get upset if you try to bring quantities back in, because quantities sound like conditions that would weaken the rule. But you don’t conclude that there are actually two tiers of utility with lexical ordering. You don’t conclude that there is actually an infinitely sharp moral gradient, some atom that moves a Planck distance (in our continuous physical universe) and sends a utility from 0 to infinity. You don’t conclude that utilities must be expressed using hyper-real numbers. Because the lower tier would simply vanish in any equation. It would never be worth the tiniest effort to recalculate for it. All decisions would be determined by the upper tier, and all thought spent thinking about the upper tier only, if the upper tier genuinely had lexical priority. As Peter Norvig once pointed out, if Asimov’s robots had strict priority for the First Law of Robotics (“A robot shall not harm a human being, nor through inaction allow a human being to come to harm”) then no robot’s behavior would ever show any sign of the other two Laws; there would always be some tiny First Law factor that would be sufficient to determine the decision. Whatever value is worth thinking about at all, must be worth trading off against all other values worth thinking about, because thought itself is a limited resource that must be traded off. When you reveal a value, you reveal a utility. I don’t say that morality should always be simple. I’ve already said that the meaning of music is more than happiness alone, more than just a pleasure center lighting up. I would rather see music composed by people than by nonsentient machine learning algorithms, so that someone should have the joy of composition; I care about the journey, as well as the destination. And I am ready to hear if you tell me that the value of music is deeper, and involves more complications, than I realize – that the valuation of this one event is more complex than I know. But that’s for one event. When it comes to multiplying by quantities and probabilities, complication is to be avoided – at least if you care more about the destination than the journey. **When you’ve reflected** on enough intuitions, **and corrected enough absurdities, you** start to **see a common denominator**, a meta-principle at work, **which one might phrase as “Shut up and multiply.”** Where music is concerned, I care about the journey. When lives are at stake, I shut up and multiply. It is more important that lives be saved, than that we conform to any particular ritual in saving them. And the optimal path to that destination is governed by laws that are simple, because they are math. **And that’s why I’m a utilitarian** – at least when I am doing something that is overwhelmingly more important than my own feelings about it – which is most of the time, because there are not many utilitarians, and many things left undone.

#### Actor-specificity: side constraints freeze action because government policies always require trade-offs since they have finite resources—the only justifiable way to resolve those conflicts is by benefiting everyone. Actor-specificity first -- different agents have different ethical obligations.

#### Lexical pre-requisite: Threats to life preclude the ability for moral actors to effectively utilize and act upon other moral theories

## CASE