

[Existential Risk / Opportunity] Singularity Management

October 2015

Contents:

- Review of "The Black Hole Case: The Injunction Against the End of the World"
- Struggling with the Ethical Limits of Expected Value Utilitarianism as Applied to Positive and Negative Singularities

Copyright 2015 Global Risk SIG. Rights, except nonexclusive multiple use, retained by authors. This publication is produced by the Global Risk Reduction Special Interest Group, a SIG within American Mensa. Content expressed here does not reflect the opinions of Mensa, which has no opinions. To join Mensa or just see what it is about, visit <http://www.us.mensa.org> . Quarterly issues of this publication are available at: <http://www.global-risk-sig.org/pub.htm> .

Review of "The Black Hole Case: The Injunction Against the End of the World" by Eric E. Johnson.

Review by Justine Jones

At the end of 2009, Eric E. Johnson, Associate Professor of Law at the University of Dakota, published a highly readable paper entitled "The Black Hole Case: The Injunction Against the End of the World" in the Tennessee Law Review (Vol 76:819). A reprint version is available at: <http://arxiv.org/abs/0912.5480> . Download the PDF.

Professor Johnson has blogged, spoken and published considerably on the subject of the dangers of particle colliders. His paper was fairly widely reported in the physics community and in the media. Here are links to some of his particle collider-related output:

"The Case of the Collider and the Great Black Hole" MIT Technology Review, physics arXiv blog, January 5, 2010, <http://www.technologyreview.com/blog/arxiv/24611>

"Law and the End of the World" Edwin Cartlidge, Physics World magazine, February 2010, pp. 12-13 <http://physicsworld.com/cws/article/indepth/41564>

Report: New Legal Solutions Needed for Experiments Like CERN's" Radio interview of Eric E. Johnson by Alex Helmick, World Radio Switzerland, February 26, 2010, <http://www.worldradio.ch/wrs/news/switzerland/new-legalsolutions-needed-for-experiments-like-ce.shtml?18077> (broken link)

“CERN on Trial: Could a Lawsuit Shut the LHC Down?”, comment and analysis, *New Scientist* 24-25 (February 17, 2010) Eric E. Johnson Page 4

What are the Chances that a Particle Collider’s Strangelets Will Destroy the Earth?” Lisa Zyga, *PsyOrg.com*, February 12, 2014, <http://phys.org/news/2014-02-chances-particle-collider-strangelets-earth.html>

<http://www.ibtimes.com/new-us-science-commission-should-look-experiments-risk-destroying-earth-1554380>

The penultimate item above refers to the RHIC at Brookhaven.

Professor Johnson also lists on his CV the following workshop and the following conference on the dangers of modern scientific research:

Judicial Review of Uncertain Risks in Scientific Research, NeTWork Workshop 2013, sponsored by FONCSI Fondation pour une culture de sécurité industrielle, Toulouse/Sorèze, France, November 22, 2013 (invited)

Science-Experiment Catastrophe and the Administrative Law Gap, Second Annual Junior Faculty Works-in-Progress Conference, Marquette University Law School, Milwaukee, Wisconsin, September 12-13, 2014 (invited)

Professor Johnson’s aim in the article reviewed here is to consider whether an injunction to stop operation of the Large Hadron Collider (LHC) at the CERN research facility in Europe could or should be able to succeed in the US courts.

A preliminary injunction is an order by the court to stop an activity pending court review of its legality. Granting such an injunction requires both hardships to the parties and the raising of serious legal questions. Hardships it can be seen would favour the plaintiffs hugely, since destruction of the world would rather outweigh the inconvenience and delay to discoveries suffered by the defendant side.(p871).

The paper provides a masterly outline of the controversy. Johnson boldly points out the inconsistencies in the arguments put by the Large Hadron Collider Safety Assessment Group (LSAG) in their 2008 ‘follow up’ to the safety report of 2003. These two reports, as he points out, do not even deal with the same arguments (p850) and the subsequent discussion in the scientific community is biased. Johnson also reports the defective response of Giddings and Mangano to Rainer Plaga’s criticisms in 2009 (849-856).

Professor Johnson clearly sees the anomalies in the ethical positions of CERN scientists: their lack of accountability due to the CERN treaty, the lack of independence of their safety reports and the pressures that would prevent scientists who did have doubts about these from speaking out. He notes the outrageous statements of John Ellis and Brian Cox from 2008 that dismiss collider criticism without consideration.(p858).

After spending almost half of the paper outlining the history of particle collider controversy and giving an excellent lay person's introduction to the scientific issues, the main criticisms, and CERN's responses, Professor Johnson moves on to legal matters (p860 onwards). He begins by exploring the precedents for an injunction to prevent harm in American law, finding that although this is not a generally acceptable principle there are examples of it happening. In particular he cites two historic cases. Firstly, *Brennan v Gellick* of 1892 (p865). This case, attempting to stop blasting of bedrock which was damaging neighbouring property, did not succeed in the courts, but only because of the plaintiff's being without 'adequate cause to fear irreparable injury.' There was an invitation to reapply if sufficient evidence actually accumulated. Johnson does not spell out the point, but it seems to me that we have plenty of evidence that a collider black hole or strangelet accident would cause 'irreparable injury'.

Secondly he cites an ultimately successful case involving railways, *Harris Stanley Coal & Land v Chesapeake & Ohio Railway Co* (ibid). This involved the rail operator trying to stop coal mining which would be likely to cause land slips and subsidence that would threaten the railroad and its operations. The eventual ruling stated that, although the risk of accident was low, the consequences were such (loss of life) that the risk was not acceptable. As Johnson says, the court's analysis is highly applicable to the black hole case. To quote the final sentences of the judgement:

'...It may be that such disaster could occur only upon a concatenation of circumstances of not too great probability and that the odds are against it. It is common experience, however, that catastrophies (sic) occur at unexpected times and in unforeseen places...A court of equity will not gamble with human life, at whatever odds, and for loss of life there is no remedy that in an equitable sense is adequate.' (p866-7)

Johnson continues by looking at the problems of jurisdiction. He earlier notes that lack of jurisdiction was the reason given for the rejection of *Wagner and Sancho's* case brought in Hawaii in 2008 (p861). CERN has immunity in all its member states; but this would not apply outside those member states.

He now brings in a most fascinating point – that in Illinois it is clear there would be jurisdiction because of CERN's 'sufficient minimum contacts' with the state. This arises because CERN uses Illinois computers and has links with Illinois universities through the LHC Computing Grid. The use of resources, including staff, who are US based by CERN for LCG, and other links such as collaboration by universities is arguably, he notes, more than enough to provide personal jurisdiction (ie i.e. the power of a court over persons) (p869). He cites *Verizon Online Services v Ralsky* (Virginia, 2002) as establishing personal jurisdiction with less contact by far.

Johnson moves on to the situation of a judge in a black hole injunction case. The problems of expert testimony are well known. Johnson explores the provision made for

US courts to ascertain the value of expert scientific testimony, the Daubert standard (https://en.wikipedia.org/wiki/Daubert_standard). (874-876). This threefold tool is difficult to apply to the collider case. Firstly it depends largely on the experts' theories being testable, falsifiable and refutable and on results being reproducible. Experimental particle physics does not subject itself to these rules on the one hand, and on the other the plaintiff's contentions could only be proved by the world ending. Secondly it asks for peer review and publication. Johnson argues that this has happened, despite rejection by peer reviewed journals of all collider critical papers, because the papers have indeed been published and been subject to discussion and scrutiny by scientists. Thirdly it cites 'general acceptance.' Johnson finds this invalid for CERN because everyone whose acceptance is valued is an interested party. The implication is that expert testimony is a dead end when it comes to a collider case.

Cost benefit analysis is also considered by Johnson, with reference to the work of Judge Posner. He critiques Posner's analysis, arguing that all Posner has shown is that people are irrational about risk taking. He then gives his own view and formulae for how life should be valued but concludes that cost benefit analysis is also of no help either in deciding a collider case (p883). He even plays devil's advocate (p883) to propose that since death is not a redressable injury under American tort law, and not everybody in the world wants to live, perhaps there is no legal downside to destroying the planet!

I would contend that filling pages with formulae may look impressive to some, but is really just window dressing. Also any figures 'plugged in' are at best highly questionable. The real core of the argument is that the loss of life would be total (including, as Johnson says, future lives and not just existing ones) and that life's value is virtually unquantifiable. I would add that neither of these gentlemen concerns themselves with the total loss of all culture, past, present, and future and (as far as we know) the only sentient life in existence. These may not be as legally important as physical life itself, and are indeed dependent on it, but surely they merit some consideration, whether you are going to look at monetary values or not.

Johnson argues that if judges fail to involve themselves and bow out they are rendering science and scientists above the law. He proceeds to the well-rehearsed arguments about the reliability of scientists. They are human and can make all kinds of mistakes (e.g. the shuttle disasters and, less known but more interesting and possibly more relevant, Castle Bravo and Castle Romeo nuclear tests in the 1950s) (p884); they can be self-interested and capable of lies and corruption.

Next Johnson urges that courts should perform a meta-analysis in deciding whether 'serious questions' are indeed raised, looking at organisational culture, group politics, and psychological context; newness of theory, reliability of data, complexity of arguments, etc, as well as the durability of pronouncements in the field. Are theories defective? Are

there mechanical, observational, or conceptual errors in the work? Are cultural, psychological, or social factors at play leading to accidental error? Are self-interest and ambition at play leading to conscious error?(p886). He gives supporting examples to illustrate all these questions. To take just one, Cambridge physicist Adrian Kent discovered mathematical mistakes in safety papers by two teams (Dar, De Rujula and Heinz and Busza, Jaffe, Wilczek and Sandweiss) relating to the RHIC (p895).

Collider critic plaintiffs would have, he notes, an embarrassment of riches if called on to support allegations of bias with evidence. (p904). ‘We must ask’, he says ‘if there is something special about particle physics that makes conflicts of interest untroubling.’

Johnson concludes that the courts must face their responsibilities rather than refusing to look at collider cases. He ends by citing some of the other scientific controversies of our day that might end up in the courts before too long, and urging readiness.

The publication of this substantial and thoroughgoing paper seemed to be quite a breakthrough for collider criticism. The professor could not be accused of going outside his field. His arguments are carefully crafted and sources quoted assiduously.

He argues against the idea that only those who are expert in a field can decide on questions related to it. Expert testimony is not to be relied on in this case, because all the experts are interested parties. (And, I would add, how would we run any sort of representative democracy or have juries sit on cases if only experts can decide?).

So, why did this groundbreaking paper and not have greater impact?

I can suggest several reasons.

Firstly, the stock of physicists is high. With the waning of belief in religious leaders they have become our high priests and conduits to mystery. They mediate exotic knowledge, knowledge that is beyond the realm of ordinary mortals, knowledge that deals with the profound. Such remote grand figures are treated as beyond questioning by ordinary mortals and above politics. So anything that challenges their authority tends to be dismissed.

Secondly, we lack idealists within the system who will risk their future careers in science by trying to call attention to the points Professor Johnson raises. Collider criticism is largely left to outsiders who are simply not considered credible by the world at large.

Thirdly Johnson is a non-scientist who is using the black hole case as a teaching tool. He has used many other bizarre examples to pique the interest of his law students (http://ericejohnson.com/exam_archive) and this can be seen as just one more. In addition he says he is agnostic on running the collider, (p886) so despite the merits of the paper and its thoroughness he is never really overtly threatening the scientific establishment at all.

So although collider critics were excited by the recognition he gave to their case, I would say that on the strength of this paper his position on colliders was not explicit enough to lead to any confrontation with the science establishment.

We should note that Professor Johnson was on leave from 2012 to 2014 and the LHC collider was shut down from April 2013 to April 2015. Naturally things have been quieter in the last few years because of this.

However, as noted at the beginning of this review Professor Johnson authored a piece about the RHIC in the International Business Times in February 2014 which demonstrated his continuing interest in collider safety.
<http://www.ibtimes.com/new-us-science-commission-should-look-experiments-risk-destroying-earth-1554380>

It would appear from the activities listed on Professor Johnson's CV and such articles as that from the International Business Times that he does in fact harbour doubts about the wisdom of running the collider. We hope that he will at some stage commit himself a little further to voicing the concerns many of us share about existential risk arising from ill-considered experiments.

Most interestingly, Johnson is due to write a piece in 2016 for the University of Illinois Law review entitled 'Agencies and Science-Experiment Catastrophes.'
(Professor Johnson's CV p3, <http://ericejohnson.com>)

It is unlikely that this will not have some relevance to collider criticism, even if it is not specifically about the LHC itself. And if it is about the LHC the fact that it is in the Illinois Law Review could be significant; Johnson argues in the 2009 paper (p868-869) that jurisdiction could be claimable there because of CERN's use of computing facilities in the state.

To conclude, The Black Hole Case is a substantial piece of work and a key text for collider critics. However, Professor Johnson's upcoming writings, both the article just mentioned and a book chapter he is working on (CV Page 4: Book Chapter 'Judicial Review of Uncertain Risks in Scientific Research' In a forthcoming book from SpringerOpen (Gilles Motet & Corinne Bieder, eds.)), will, we, hope, augment this with more explicit and controversial criticism that truly confronts the science establishment.

Struggling with the Ethical Limits of Expected Value

Utilitarianism as Applied to Positive and Negative Singularities

by James Blodgett

How do we choose where to steer potential singularities? How do we address negative singularities, i.e. existential risks? Before we steer anywhere we need a direction, and we need a compass to point that direction. The greatest good for the greatest number, the goal of utilitarianism, sounds like it might be that compass. But we need to look deeper.

When there are many different possibilities, it sounds like expected value is the way to evaluate them. Expected value is the sum of the value of the various outcomes of a choice, each multiplied by its probability. Ideally, choices are made by computing the expected value of each of the menu of choices one might make, and implementing the choice that has the highest expected value. However, this is not always easy to do. When there are many different possibilities of uncertain probability and uncertain value, the math of expected value becomes a subjective estimate. When we attempt to steer humanity as a whole, the ultimate direction steered becomes a consensus of the many hands on the steering wheel, so we have a conflict of subjective estimates from many viewpoints. However, this situation is not as bad as it sounds. We are guaranteed a direction since not to decide is to decide not to decide. Also, a lesson of democracy is that the consensus of diverse viewpoints is often better than the viewpoints of single individuals. Also, this picture of a global consensus gives us personally something to do that can contribute, which is to contribute to the discussion. Our ideas have some chance of becoming memes that contribute to the debate. People with power often make decisions in the context of that debate.

An advantage of seeing expected value estimates as subjective is that at least we can always make them. If we try hard and honestly we can sometimes make them well. If we think deeply about how we made them we can also estimate their reliability. Precision is rare, but there is often enough reliability to point a direction.

One input to our estimates is to contemplate known problems within the math of expected value.

Since our business is existential risk and existential opportunity, and that involves lots of human lives, it is convenient to denominate the value we put into our expected value equation in terms of human lives. For example, suppose a group of mad scientists (CERN would be a good example) risks destroying the world via creation of black holes or strangelets or vacuum transitions with a probability of $P = 0.0001$. {Since we are

using CERN as an example, we should note that they aren't precisely mad scientists. They don't cackle evilly. Most got into science for reasons that are noble without realizing downsides, and most justify continuing despite downsides by accurately concluding that downsides are fairly improbable, since downsides require several levels of speculative science to actualize. However, most of cutting-edge physics is speculative, so speculative science may indeed actualize, and many CERN scientists will indeed be mad (mad at me) if they read this analogy picturing them as mad scientists} ... Resuming, if the probability of disaster is $P=0.0001$, the expected value (negative in this case) in terms of human lives is ... {assuming we believe the estimate of $P=0.0001$, which is a very rough estimate. Collider critics have published higher estimates, including the estimate that $P=0.5$ (based on humanity's lack of knowledge of the underlying physics, a lack that is somewhat accurate since CERN explores the unknown. If we are truly ignorant about which of two outcomes will occur, as in flipping a coin, it is appropriate to estimate a probability of $P=0.5$ for each. However, I think it rather silly to claim total ignorance of whether disaster will occur or not in this case, since it is known that the physics that predicts disaster is speculative, and one of many speculations.) Collider advocates have offered estimates that are lower than $P=0.0001$, and even developed a new safety consideration, but those lower estimates are difficult to sustain because Ord et al point out that any such very low probability estimate is overwhelmed by the not-as-low probability that the theory on which the estimate is based is wrong, not unlikely when exploring unknown physics. [Toby Ord, Rafaela Hillerbrand, and Anders Sandberg, "Probing the Improbable: Methodological Challenges for Risks with Low Probabilities and High Stakes," *Journal of Risk Research* 13(2) (2010) pp. 191–205. See <http://arxiv.org/abs/0810.5515> . Ord et al use CERN as an example.] Real theorists have demonstrated that the theories underlying CERN's safety considerations may be wrong by finding plausible if speculative theory that circumvents most safety considerations. Presumably realizing that they could not support the very low estimates appropriate for safety when Earth is at risk, and concerned about public reaction, CERN adopted the official estimate that $P=0$, established by fiat. $P=0$ is precisely equivalent to that famous statement of scientific hubris, "Nothing can possibly go wrong." The Daily Show appropriately made fun of this range of estimates, interviewing an advocate of $P=0.5$ and an advocate of $P=0$.

This range of estimates is not just a matter for humor. It is symptomatic of the difficulty and the fragility of subjective probability estimates. As we see here, they are heavily influenced by personal or organizational interest conflicting with the interest a scientist or a planner has in accurately predicting the future. Subjective probability estimates are also heavily influenced by the mental model in the mind of the person doing the prediction. Small changes in that model can dramatically change even the most objective probability estimate. As an example, an estimate of the limits of the lifetime of things without a definitive lifetime can be estimated from the amount of time they have already existed. For example, incorporated businesses don't die on an actuarial curve

like humans, but they do succeed or fail. If we assume that it is unlikely that they are either in the first or the last one percent of their existence, this assumption predicts a longer life for those that have already existed for a while. Studies have shown that this prediction is usually born out; those that have been around for a long time are more likely to stay around. So consider the application of this to existential risk. The human species has been around for many thousands of years. This would seem to lead to the assurance that we are likely to be around for at least a few more centuries. However, that depends on how one defines our species. For example, does Neanderthal count? There have been natural mass extinctions in Earth's history, but the natural version seems to happen infrequently. Right now the biggest threat seems to be misapplications of a new thing, human technology. Therefore, it might be relevant to consider the length of time that humans with the technology to destroy our species have existed. Humans with this technology have existed for only about fifty years. (It was a few years after Hiroshima before we had enough hydrogen bombs to really do the job.) Applying the implied math, the simple fact that we are a species that has had existential risk technology for a fairly short time suggests that the probability that we will exist for much longer is fairly low. Willard Wells uses a sophisticated version of this math to calculate our odds in [Willard Wells, *Apocalypse When?: Calculating How Long the Human Race Will Survive*, Springer, 2009.] The sensitivity of these estimates to assumptions is illustrated by Wells himself, who is about to publish a new book with a lower estimate of our chances using a refinement of his original method.

} (the end of the parenthetical diversion started three paragraphs ago.) We should talk to the Guinness World Records people about this number of parenthetical diversions, diversions that are relevant, but that may obscure the point we are developing, which is that: ... given $P=0.0001$ of an existential risk, the expected value in terms of human life is P times the population of Earth (plus our expectation of future population, but this is difficult to estimate and somewhat of a distraction) or 0.0001×7 billion = 700,000 human lives. Hitler did worse, but this is still not a good thing. Hitler did worse not only by killing more, but by actually killing them. CERN has the slightly better moral position of only risking a small chance of killing-- but it is a chance not only of killing 700,000, but of killing 7,000,000,000. The expected value of that chance is properly reduced by weighting by probability. But what is to be said of the morality of that chance?

It is relevant that expected value works in the positive direction too, and sometimes gives larger absolute values in that direction. Some positive singularities can enable a large number of human lives. If so, I call the potential for such an event an existential opportunity, which is like an existential risk (a risk to human existence) but goes in the other direction, enabling a transcendent increase in human numbers rather than our extinction. For example, Lewis estimates that there is enough material in our asteroid

belt to build habitats for 10,000,000,000,000,000 people. [John Lewis, *Mining the Sky: Untold Riches from the Asteroids, Comets, and Planets*, Perseus Publishing, 1997, pg. 194.] This estimate is rough for several reasons. First, it is based only on the amount of iron needed for the structure. However, the asteroid belt has lots of frozen water (which is partly oxygen) in ice asteroids, and all of the other elements we have on Earth, so there may be enough to fill all of those structures with biospheres. A larger problem is sustainability. It is unlikely that an asteroid civilization with those humongous resources would conserve them strictly, and they would have lots of reasons to use lots of those resources for purposes other than biospheres. For example, if they want to travel almost anywhere, they would have to throw some mass away as reaction mass. (Rockets move by throwing exhaust in one direction so as to move in the other. Rockets we think of as standard accelerate reaction mass by forcing hot gas through a nozzle, but mass can also be accelerated by magnetic fields.) I say "if they want to travel almost anywhere" since O'Neill, who designed attractive rotating cylindrical space habitats, suggested a way to visit neighboring habitats without using reaction mass. If they are rotating in the right plane, one simply releases a shuttle when it is aimed right. The shuttle will retain the speed of the rotation and travel to the other habitat. Many of O'Neill's habitats will have a twin rotating nearby in precisely the right plane because of another of O'Neill's design tricks, a method to spin up massive habitats without using reaction mass. O'Neill's trick is to link two habitats, and use a small motor for a long time to spin them up so they are rotating in opposite directions. This can be done without reaction mass since each is pushing the other. This also results in an assembly of two habitats with a collective angular momentum of zero. This helps to keep their mirrors pointed at the sun as they move in orbit despite the gyroscopic effect, since the gyroscopic effect of one would counter that of the other. The joint assembly can have a small rotation that keeps mirrors pointed despite their movement in orbit. The shuttle trick can work approximately for nearby habitats that are not linked; O'Neill estimated that a population similar to that of Earth could visit among themselves in this way.

Despite the fact that Lewis's estimate may be difficult to achieve in practice, it still suggests that a humongous number of human lives might be sustained in space even if we make large allowances for loss of mass for reaction mass, leakage, etc. Also, Lewis's estimate involves only the asteroid belt. There is enough mass in space as a whole to permit expansion of his estimate by many many orders of magnitude. When calculating expected value, staying in our solar system increases the other component of the calculation, probability, since O'Neill's ideas seem fairly plausible, and Metzger's ideas for making them happen, discussed here in past issues, make them more plausible. However, settling much of the universe is not totally implausible despite the light speed limit, since given appropriate technology it could be done via many seed ships that reconstruct infrastructure and life, including human life, on arrival at a suitable concentration of mass.

The possibility of existential opportunity gives us one reason that CERN may be right. Physics just might come up with some miracle energy source or space drive that enables some marvelous singularity such as industrialization of space, and CERN research just might be key to the discovery. The relevant probabilities are difficult to estimate. If some great discovery is waiting for us, is it likely that CERN research is the only path to that discovery? Plug in the right numbers, and the expected value of a positive singularity via CERN (and no one else) might seem to outweigh the negative expected value of extinction via CERN (and no one else, or at least no one else right away, at least giving us a few more lives to add to that utilitarian greatest number.) But even if so, does this work morally? Is it right to risk the lives of currently living people to enable other probabilistic future lives? Trolley problems used in ethical philosophy can be ambiguous. Most people would push a switch to divert an out-of-control trolley from a track where ten men are working, to a track containing only one man. However, most people would not push a fat man onto the track, using his mass to stop the trolley, although the numbers are the same.

Expected value has problems at the extremes. One of these problems is well known. Expected value often uses value denominated in money. However, a small chance of receiving a large amount of money may not be worth its expected value because the utility of a small unit of money varies as the amount possessed increases. A dollar to a starving man may represent his next meal, whereas a dollar to a rich man may be less than the amount he will give the waiter as a tip. (In economics jargon, the marginal utility of money declines.) Thus, if we are poor, the good outcome of a business decision with the potential to make us billionaires is not worth its expected value denominated in money since money is not worth as much to the billionaires we would become if it works. A solution is to denominate in utility, although the amount of utility is harder to estimate. Here we are denominating in human lives. Morally, one life should be as valuable as another. Nevertheless, we tend to make a discount similar to that of marginal utility to the value of human lives in large populations, because we are worried about population problems, which in the worst case can be existential risks. That discount did not apply when settling the American continents when they were sparsely populated, and it would seem not to apply when we are at the early stages of settling the vast resources of space if our technology gets to the point where it can facilitate access to and use of those resources. We also tend to make a discount to the value of other lives because they are not our own, and most are not even acquaintances. However, this discount should not apply when we are considering public policy from the point of view of what is best for the public.

Denominating expected value in human lives is part of utilitarianism, which aims for the greatest good for the greatest number, since the number of human lives attainable is precisely that greatest number. Utilitarianism seems ethical, but that ethics can be debated when utilitarian considerations contradict another form of ethics known as

deontology. Deontology involves lists of actions that are moral and actions that are not. For example, "Thou shalt not kill." A crude utilitarianism might approve killing someone if that killing saves many others via organ transplants. I think most of us would reject that version. But suppose those on the downside of such a transaction were not killed outright, but only put at risk, perhaps a very small risk? For example bone marrow for transplantation, valuable for reestablishing the immune system after chemotherapy, can be taken from live donors, at a small risk to the donor. Some people volunteer to make such donations. Deontology provides a reason to question balancing a chance of eliminating existing real lives with a chance of enabling probabilistic future lives. Different ethical systems might call this one differently.

The ethics of these issues, the estimates of probability, and the estimates of utility are often difficult to specify precisely. However, we are constrained to solve these problems somehow, since "not to decide is to decide not to decide." It helps to make the best use that we can of the information and the philosophy that we do have available, so that we can get closer to making the best decisions that limited humans, or indeed any other limited entity or system, can make. We can help by contributing to thought on these issues.

One way to contribute to thought on these issues is to write an article for this publication.