

Johann Frick  
**Overview of my Research**

My research spans a range of topics in moral and political philosophy, practical reason, and applied ethics. In this document, I describe my publications and ongoing research on (1) the ethics of risk, (2) population ethics, (3) interpersonal normativity, (4) the morality of nationalism and the ethics of immigration, and (5) structural rationality and self-knowledge. I also lay out some of the cross-cutting thematic connections between my work in these different areas. Section 6 contains my own assessment of what constitutes my best work so far, while Section 7 describes some of the professional honors I have received and the impact that my research has had to date. (A complete list of works that discuss or cite my research is contained in the Appendix, beginning on page 23).

**Disclaimer:** My regular tenure review at Princeton University was scheduled for next academic year, and this is the date my publication strategy has been geared towards. Since I only want to submit work of the highest level of polish for your review, working papers (marked with an asterisk) that are not yet published or under review are not publicly available on my website. However, in order to help you to better situate the papers that are included in my file in the context of my larger research agenda, and to give you an idea of some of the other projects that I currently have underway, the overview also contains brief descriptions of some of my other papers in progress.

## **1. Ethics of Risk**

Moral philosophers, to date, have paid insufficient attention to questions about risk and uncertainty. In discussing the morality of imposing harms and conferring benefits on other persons, philosophers have focused overwhelmingly on cases with *certain* outcomes, where the agent knows both (i) *who* will be affected by her action or policy and (ii) *how* they will be affected, i.e. whether they will be benefited or harmed, and by how much. In reality, we often lack knowledge of (i) or (ii) or both. In many cases, actions that have the potential to benefit a given person also come with a downside risk of harm, to this individual or to others. What ethical principles ought to govern our thinking about such cases?

In my article “**Contractualism and Social Risk**” (*Philosophy & Public Affairs*, 2015), I focus on risky actions or policies that affect large groups of people – what I call the imposition of “social risks”. In such cases, as they occur routinely in public or health policy, statistical knowledge often allows us to predict, with a high level of confidence, the overall *pattern* of outcomes across individuals that an action or policy will produce. At the individual level, by contrast, we cannot know in advance which persons will be benefited or harmed. In such cases, I say, both the agent and the patients find themselves behind a “natural veil of ignorance” that precludes them from knowing the individual-level effects of the action.

Previous scholars had argued that cases of social risk pose insuperable problems for many non-consequentialist moral views, and in particular for moral contractualism. One of the defining characteristics of much contemporary nonconsequentialist thought is skepticism about *interpersonal aggregation*. Leading nonconsequentialists have put forward “competing claims”

models of moral rightness, according to which morality requires us to determine, by a series of pairwise comparisons, that action or policy which satisfies the strongest individual claim (or generates the weakest individual complaint). The fullest development of this anti-aggregative approach is the contractualist moral theory defended by T.M. Scanlon in *What We Owe to Each Other*.

Many cases of social risk are *prima facie* problematic for the contractualist, because they are characterized by the following four features: (1) The risky action will affect a large number of individuals. Because of this, it is virtually certain that *some* people will end up being burdened by it. (2) The individual losses to those who are burdened are considerably greater than the individual gains for those who are benefited. (3) The action-type in question is rare, or rarely affects the same people twice; as a result, we cannot assume that over time almost everyone will benefit from a principle that permits actions of this type to be performed. And yet, despite (1), (2), and (3), it seems that (4) the risky action is intuitively permissible.

In my paper, I argue that the problem of social risk forces the contractualist to confront a question about his theory that has not received a fully satisfactory answer to date: Are an individual's personal reasons for rejecting a principle permitting the imposition of a given social risk a function of the *prospect* that the risky action gives to each person *ex ante*, allowing us to discount both benefits and harms by their improbability of occurring? Or is justifiability to each person a function of the action's outcomes *ex post*? What matters morally isn't the *ex ante* likelihood of any *given* individual's being benefited or harmed by the action, but rather the near certainty that, *ex post*, *some* persons will turn out to have been harmed by the action while others are benefited.

The contractualist faces an apparent dilemma: if he rejects the *ex ante* view of justification and appeals instead to an *ex post* view, he seems committed to the intuitively unappealing conclusion that most instances of social risk are unjustifiable. On the other hand, if he embraces an *ex ante* view, Scanlon fears that this will move his contractualist theory too close to those aggregative views that he wished to escape from in the first place.

In my paper, I elaborate a version of contractualism, what I call *stage-wise ex ante contractualism*, that solves the problem of social risk without resorting to interpersonal aggregation. Contractualists, I argue, can "count the numbers without aggregating". My solution, however, comes at a price: Contractualism can no longer claim to be the correct account of what it is for an action to be morally wrong, all things considered. Instead, contractualism can at most claim to capture one *class* of moral considerations that *contribute* to making actions right or wrong, all things considered.

A second paper on risk, "**Uncertainty and Justifiability to Each Person**" in Eyal, Hurst, Norheim and Wikler (eds.), *Inequalities in Health: Concepts, Measures, and Ethics* (New York: Oxford University Press, 2013), concerns the intersection of questions of risk and distributive ethics. It critiques the arguments of Marc Fleurbaey and Alex Voorhoeve in their important article "Decide as You Would with Full Information! An Argument against *ex ante* Pareto" (same volume). According to the *ex ante* Pareto principle, which is widely appealed to in welfare economics and in the literature on health inequalities, if an alternative has higher expected utility for every person than every other alternative, then this alternative should be chosen. Fleurbaey and

Voorhoeve claim that ex ante Pareto is not just inconsistent with impersonal distributive principles such as telic egalitarianism or prioritarianism; it also violates a broadly contractualist criterion of moral rightness, according to which an action is right if and only if it is justifiable to each person. I argue that Fleurbaey and Voorhoeve's main argument against ex ante Pareto fails. The Principle of Full Information, on which it is based, is either unsound, or else cannot settle the debate between defenders and opponents of ex ante Pareto. I conclude that while outcome-based moral principles, such as telic egalitarianism or prioritarianism, do indeed militate against ex ante Pareto, there is no objection to this principle from the vantage point of justifiability to each person.

Although the core of my work on the ethics of risk is theoretical, it has important practical upshots. One of these concerns the so-called problem of "identified" vs. "statistical" lives. The identified vs. statistical lives issue is of great relevance to many debates in practical ethics, for instance the debate between proponents of *treatment* and *prevention* in the battle against HIV/AIDS: For years, debates about the best way to combat the AIDS pandemic have pitted proponents of scaling up antiretroviral treatment for people suffering from AIDS against those advocating for more cost-effective prevention measures. In an influential recent article, Dan Brock and Daniel Wikler argue that there is no sound moral basis to privilege the saving of identified lives through antiretroviral treatment, if preventive methods could save more (statistical) lives.

In my paper "**Treatment versus Prevention in the Fight Against HIV/AIDS and the Problem of Identified versus Statistical Lives**", in Cohen, Daniels, and Eyal (eds.), *Identified versus Statistical Lives: An Interdisciplinary Perspective* (New York: Oxford University Press, 2015), I take issue with Brock and Wikler's argument. In so doing, I develop a novel account of how the choice between "treatment" and "prevention" turns, in part, on the question whether we have reasons of fairness to want risks of harm to be spread more "thinly" across a larger number of people rather than being concentrated in a few, even if the expected number of lives lost is greater in the former scenario. The paper concludes with a postscript on "treatment-as-prevention" (TasP), a new avenue of AIDS research which stresses the preventive benefits of early antiretroviral treatment. I argue that, despite its medical promise, TasP does not transcend the ethical dichotomy between treatment and prevention explored in this paper.

While my published papers on risk deal with a variety of *forward-looking* ethical questions, concerning the permissibility of imposing risks on others and their equitable distribution, in my more recent research, I turn my attention to a range of *backward-looking* ethical issues, for instance questions concerning our moral responsibility or culpability for outcomes that were produced by risky (or 'chancy') causal processes, as well as questions surrounding the appropriate compensation of people who have been harmed by such processes.

As a prolegomenon for this work, I undertook a study of the metaphysics literature on probabilistic causation. From this work emerged my next paper, "**Chancy Causation and the Problem of Aggregate Effects**" (R&R at *Philosophical Studies*) which critiques David Lewis's prominent account of probabilistic causation. According to Lewis, an actual event  $c$  is a cause of an actual event  $e$  if and only if (i)  $e$  causally depends on  $c$ , or (ii) there is a chain of events  $d_1, d_2, \dots, d_n$ , such that  $d_1$  causally depends on  $c$ ,  $d_2$  causally depends on  $d_1$ , ..., and  $e$  causally depends on  $d_n$ , where  $e$  causally depends on  $c$  if and only if, had  $c$  not occurred,  $e$ 's probability of occurring would have been *smaller by a large factor* than its actual chance of occurring.

It is a direct corollary of Lewis's causal dependence account that an event  $c$  that only *slightly* raises  $e$ 's probability of occurring cannot be a cause of  $e$ . While not overly troubling at first blush, my paper argues that this feature of Lewis's view has implications that are hard to defend. The first of these problems is that Lewis's theory implies, implausibly, that an event  $c$  may cause an "aggregate event"  $E$  without being a cause of any of the "constituent events"  $e_1, e_2, \dots, e_n$  that together *make up* the aggregate event. This is the Problem of Aggregate Effects. Lewis' view also implies, again implausibly, that a group of events  $c_1, c_2, \dots, c_n$  can together cause an effect  $e$ , without *any* of  $c_1, c_2, \dots, c_n$  being a cause of  $e$ . This is the Problem of Aggregate Causes. Finally, it implies that one and the same event  $c$  can fail to be a cause of  $e$ 's occurring yet also be a cause of  $f$ 's *not* occurring, even when  $f$  just is "the non-occurrence of  $e$ ". This is the Problem of Causation and Prevention. These three problems, I claim, seriously undermine the attractiveness of a causal dependence approach to probabilistic causation, lending support instead to an account in terms of causal *contribution*.

My working paper \***"Risk, Responsibility, and Aggregate Effects"** turns from metaphysical questions about causation to moral questions about responsibility and culpability in the production of aggregate effects. I show that moral responsibility at the level of aggregate and of constituent events can come apart in various ways. For instance, I argue that (1) in virtue of having performed some action  $\phi$ , an agent can have a high degree of moral responsibility for some *aggregate* effect  $E$ , while having only a low degree of responsibility for any of the constituent events that together *make up*  $E$ .<sup>1</sup> In such cases, an agent (or group of agents) can be significantly culpable, given their responsibility for bringing about a result that is very bad in the aggregate, without significantly *wronging* anyone affected by his action. But likewise, I show that (2) when the actions of many agents, acting independently of each other, together produce an aggregate outcome for an individual that significantly harms this person, this person can have suffered a significant *wrong*, without anyone's behavior having been significantly culpable – indeed without any individual agent having significantly *wronged* the patient. I argue that like "to harm" but unlike, say, "to love" or "to believe", the transitive gradable verb "to wrong" is what I call a *cumulative* verb, i.e. a verb of which the following is true: For it to be the case that a patient  $P$  is  $\phi$ -ed to a high degree, there need not exist any individual agent  $A$  who  $\phi$ s  $P$  to a high degree. It is enough for many agents to  $\phi$   $P$  to a low degree.

---

<sup>1</sup> Consider the following illustration: A construction company is charged with building a new highway. It has a choice between two route designs: a straight and monotonous design, or a more varied design. The company knows from private safety studies that, all else equal, a driver's risk of falling asleep and having a serious accident is *slightly* higher for monotonous routes (11/10,000 on a given journey) than for more varied routes (10/10,000 on a given journey). Nonetheless, the company picks the monotonous design, to save a trivial amount of money. As was foreseeable, the number of accidents on the new highway involving drivers who fall asleep at the wheel is appreciably higher than on comparable highways with a less monotonous design. Ask yourself "who is morally responsible for the unusually high *level* of accidents on this road?" and the answer seems clear: it is the road construction company, due to its ill-advised choice of road design. By contrast, ask yourself "what is the company's degree of moral responsibility for any *particular* accident in which a driver falls asleep?" and it seems the answer is: only a *very minor* one. After all, a given driver is only at a *slightly* greater risk of falling asleep under the monotonous design; and for any driver who does fall asleep and have an accident, other factors (not having gotten enough sleep the night before; having taken soporifics before going on the road, etc.) will have played a far more important role than the monotony of the road design.

My paper shows that these theoretical results have important implications for our thinking about many issues in ethics and public policy, among them (i) questions about liability for risky harms, as they arise, for instance, in tort law; (ii) questions about social responsibility for health inequalities between socioeconomic groups, as they are studied by bioethicists and epidemiologists; and (iii) questions about structural injustices and individual culpability, which feature prominently in contemporary social justice discourse. For instance, I argue that attention to (1) can help resolve a long-standing but unproductive debate between ‘liberals’ and ‘conservatives’ about individual vs social responsibility for health: we can recognize that the socioeconomic gradient in society explains much of the aggregate-level health inequalities between economically better- and worse-off *groups*, while still acknowledging that individuals may have the lion’s share of responsibility for their private health outcomes (especially as far as so-called ‘life-style diseases’ are concerned). Societal responsibility for health inequalities between social *groups* does not “crowd out” individual responsibility for health. Likewise, if (2) is true, this supports the contention, by Iris Marion Young and others, that injustice does not require a ‘bad actor’. A person can be seriously wronged, or be the victim of serious injustice, without there being any individual agent who has seriously wronged that person, or treated her seriously unjustly.

A further working paper, **\*“A Puzzle About Risk and Compensation,”** discusses principles for the compensation of harms that were caused by impermissible risky actions, as they feature in tort law. I argue for an impossibility result that holds for all systems of compensation in which the victim of a tortious harm is compensated directly by the tortfeasor. In any such system, I argue, it will be impossible to simultaneously satisfy all three of the following intuitively plausible desiderata. We can at best satisfy any two:

- (1) Absent special considerations (e.g. punitive damages, etc.) the total compensation that an agent pays should be equal to the total harm that this agent is causally responsible for.
- (2) The amount of compensation that a victim receives from the agent should be proportionate to the degree to which her interests were set back by the agent’s action.
- (3) Victims of an impermissible risky action should receive greater compensation from the agent than non-victims (i.e. persons who were exposed to a risk of harm but didn’t suffer any harm).

Violations of each of these desiderata lead to worries about horizontal equity across different cases. If desideratum (1) is violated while the other two are satisfied, tortfeasors in different cases will have to pay unequal amounts of compensation, despite having caused the same amount of total harm. If desideratum (2) is violated while (1) and (3) are satisfied, victims will receive different amounts of compensation in different cases, despite their interests having been set back to the same degree. If desideratum (3) is violated while (1) and (2) are satisfied, there will be cases in which victims and non-victims receive equal amounts of compensation.

Finally, I am also working on a **book project on the ethics of risk**: There are currently very few book-length treatments of the ethics of risk by a philosopher.<sup>2</sup> Peter Momtchiloff of Oxford University Press has solicited me to write such a book, and I am currently working on a book

---

<sup>2</sup> John Oberdiek’s *Imposing Risk: A Normative Framework* (OUP, 2017) is the notable exception.

proposal which I plan to submit by the end of this summer. The book (working title *Morality Behind a Natural Veil of Ignorance: Risks, Rights, and Responsibility*) will draw together the results of my research into the ethics of risk, and spell out its implications for a wide range of applied issues in practical ethics and public policy, in tort law, etc.

## 2. Population Ethics

My second major area of research over the past few years has been population ethics, the topic of my doctoral dissertation *'Making People Happy, Not Making Happy People': A Defense of the Asymmetry Intuition in Population Ethics* which won Harvard University's Emily and Charles Carrier Prize "for an outstanding doctoral thesis on a subject in social, political, or moral philosophy" in 2015.

Since the seminal contributions of Derek Parfit and Jan Narveson in the 1970s and 80s, the field of population ethics has largely revolved around four major problems, to which various theorists have attempted to provide solutions. These are the Procreation Asymmetry; the Non-Identity Problem; the Repugnant Conclusion; and the Mere Addition Paradox. Oftentimes, a proposal that appears promising in response to one of these four problems has disastrous implications elsewhere. For this reason, I have waited to publish my work in this area until I felt confident that I had a viable and coherent account of the entire domain. I am now in the process of completing work on a series of six articles that offer novel solutions to all four problems. Two are published, two are currently going through the review process at *Philosophy & Public Affairs* and at *Noûs*, and the final two are committed to venues where they will be submitted for publication within the next few months.

The starting point of this strand of my work is to defend the so-called Procreation Asymmetry, a widely held moral intuition according to which, while there is a strong moral reason not to create a life that we know will not be worth living, there is no moral reason to create a life *just because* it would be worth living. This intuition is presupposed by much common-sense moral thought about procreation and about social policies that will foreseeably impact the size of future generations. Moreover, denying the intuition opens the door to other notorious theoretical problems in population ethics, above all Parfit's Repugnant Conclusion. Nonetheless, the Procreation Asymmetry is only defended by a minority of philosophers working in the field of population ethics. This is because any attempt to give a systematic account of the intuition soon runs into theoretical problems which many have considered insuperable.

My first article and the flagship paper of my project, "**Conditional Reasons and the Procreation Asymmetry**" (currently R&R at *Philosophy & Public Affairs*) investigates one such difficulty. I show that one reason why previous attempts to justify the Asymmetry have failed is that they have counterintuitive implications in "non-identity" cases – that is, cases in which there is a choice between causing a better-off or a less well-off individual to exist. These existing attempts to vindicate the Procreation Asymmetry all pursue a common strategy: While totalist utilitarianism, which commits us to rejecting the Asymmetry, holds that the potential well-being of *all* possible people is reason-giving, these proposals maintain that only the potential wellbeing of *some* possible people is reason-giving. They single out some possible people according to their temporal

location (e.g. *presently existing* individuals) or their modal status (e.g. individuals who exist *independently* of the procreative choice, or individuals who are ever *actual*) and maintain that only the potential well-being of *these* possible people is reason-giving. I show that all such attempts to defend the Asymmetry fail, and fail for the same reason: the very feature that allows these accounts to capture one or both conjuncts of the Asymmetry precludes them from respecting most people's intuition about non-identity cases, namely that, all else equal, it would be wrong to cause the less well-off rather than the better-off individual to exist.

Extant attempts to vindicate the Asymmetry fail, because they focus on the wrong *aspect* of totalism, while letting its crucial assumption go unchallenged. The Asymmetry must be explained, not by challenging totalism on *whose* well-being matters, but on *how* people's well-being matters, i.e. on its account of the reason-giving force of wellbeing.

Totalist utilitarianism is a teleological moral view. According to the teleologist, we have *unconditional* reasons to maximally promote that which is good, and to dispromote that which is bad, because this is what produces the best state of affairs. Hence, there is no deep moral distinction between 'making people happy' and 'making happy people': both lead to better states of affairs, by making it the case that that which is good (namely wellbeing) is more widely instantiated. By contrast, I argue that the key which unlocks the philosophical puzzle of the Procreation Asymmetry is recognizing that our reasons to confer well-being on a person S are not unconditional reasons to promote well-being, but rather *conditional* reasons, which have their source in S, a being with moral status. If this is correct, it explains why there is no unconditional moral reason to create a person, just on account of the well-being we would thereby confer on her, since any welfare-related reasons to confer well-being on a person are conditional on her existence. The morality of procreation is instructively parallel to that of promise-making: Just as it is wrong to make a promise that one cannot keep, so it is wrong to cause a person to exist if one cannot ensure that she will have a life worth living. Likewise, just as we don't have reasons to make promises, just because we would keep them, so too the mere fact that a possible person's life would be worth-living does not give us moral reason to create her. However, *conditional* on it being the case that one will create a new child, one has moral reason to create the individual whose life will foreseeably go best. My view thus allows me to give the correct answer in Non-Identity Case while simultaneously upholding the Procreation Asymmetry.

While "Conditional Reasons and the Procreation Asymmetry" states my positive account of the Procreation Asymmetry, two further papers aim to defang two of the most common and potent objections to this intuition. The first of these is "**On the Survival of Humanity**" (*Canadian Journal of Philosophy*, 2017), since reprinted in the Routledge Volume *Ethics and Future Generations* (ed. Rahul Kumar). The paper asks: What moral reasons, if any, do we have to ensure the long-term survival of humanity? Most responses in the literature are premised on denying the Asymmetry intuition. We have moral reasons to ensure that humanity survives for as long as possible, it is said, because, all things equal, this will maximize the number of worthwhile lives that are lived. If embracing the Asymmetry left us with no basis for being morally concerned about the long-term survival of humanity, this, to my mind, would be tantamount to a *reductio* of the intuition. In my chapter, I explore what resources there are for an alternative account. I argue that we may have moral reasons to care about the long-term survival of humanity because seeking to sustain humanity into the future is an appropriate response to the *final* value of humanity itself. On

this account, the moral reasons we have for ensuring the survival of humanity do not also constitute arguments for creating additional lives in the present. In addition, the paper contains an original critique of John Broome’s influential “greediness objection” against the Intuition of Neutrality (the axiological counterpart of the Procreation Asymmetry).

In the third article, \***“The Procreation Asymmetry and the Specter of Antinatalism”** (to appear in a Festschrift in memory of Derek Parfit, edited by T. Campbell, J. McMahan, and K. Ramakrishnan), I explore some implications that my defense of the Asymmetry has for questions of individual procreation. I discuss and reject an argument due to David Benatar (2006), according to which accepting the Asymmetry commits us to recognizing a prima facie obligation against having any children. More challenging is the following problem, which I term “the problem of procreative risk”: In every act of procreation, there is some small risk of creating a person whose life is not worth living. Does this empirical fact, together with the Procreation Asymmetry, commit us to the view that procreating is a moral risk that is rarely justifiable? Consider: if we procreate then, in the best case, we bring about an outcome (namely that in which a new happy person is created) which, according to the Asymmetry, we have no positive moral reason for bringing about. But in the worst case, we bring about an outcome that we have strong moral reason to avoid. Drawing some of my recent work on the ethics of risk, I argue that even this second challenge can ultimately be met; affirming the Asymmetry does not, in fact, commit us to antinatalism. But the argument for this conclusion is subtle, and requires us to contend with hitherto unexplored questions in the morality of risk imposition in procreative contexts.

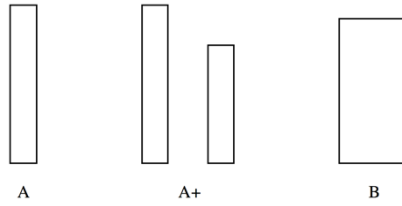
The fourth paper in the series, **“Context-Dependent Betterness and the Mere Addition Paradox”** (currently under review at the *Journal of Moral Philosophy*, as part of a special issue on population ethics edited by Tina Rulli and myself), deals with the final two of Parfit’s famous puzzles, the Repugnant Conclusion and the Mere Addition Paradox.

The simplest arguments for the Repugnant Conclusion are “spectrum” arguments, as illustrated by the following figure: Each block represents a different possible history of the world, with the width of the block representing the number of people who



ever live and the height of the block their quality of life. We are asked to assume that B is better than A, because, while the people in B are less happy than the people in A, there are sufficiently more of them to make B all things considered better than A. Iterated application of this step, plus the transitivity of “all-things-considered better than”, lead us to the repugnant conclusion that Z, an extremely large population made up of people whose lives are barely worth living, is better than A, a population of 10 billion people with very happy lives. Embracing the Procreation Asymmetry (or rather its axiological counterpart, the Intuition of Neutrality, according to which, while creating a new life that is not worth living makes the world worse, all else equal, creating a new life that is worth living does not make the world better, all else equal) provides a straightforward way of resisting such simple spectrum arguments for the Repugnant Conclusion. We simply deny the first step in the argument, the claim that B is all-things-considered better than A, on account of containing a sufficiently greater number of happy lives.





Parfit's Mere Addition Paradox, illustrated by the figure to the left, requires a special treatment, however. In *Reasons and Persons*, Parfit very convincingly argues that each of the following three pairwise comparisons is extremely plausible: (1) A is better than B; (2) B is better than A+; (3) A+ is not worse than A. These three claims, however, appear to form an inconsistent triad. Lest we give up on the transitivity of "all-things-considered better than", it seems, we must abandon one of (1), (2), or (3). The claim that we have least reason against giving up, Parfit thinks, is claim (1). However, if A+ is not worse than A, and B is better than A+, then B must be *better* than A. This argument can now be iterated, to again take us to the Repugnant Conclusion. Embracing the Intuition of Neutrality does not help us to resist *this* argument, since none of its premises rely on denying the Intuition of Neutrality.

In my paper, I argue that the Mere Addition Paradox trades on an ambiguity about the context of choice. f

- (1\*) A is better than B, when the option set is  $Q = \{A, B\}$ ;
- (2\*) B is better than A+, when the option set is  $R = \{A+, B\}$ ;
- (3\*) A+ is not worse than A, when the option set is  $T = \{A, A+\}$ .

The air of paradox, I think, arises from the assumption that the pairwise judgments (1) - (3), which I have argued are true in *two-possible* cases, carry over to a *three-possible* case. It would then be true that

- (1') A is better than B, when the option set is  $S = \{A, A+, B\}$ ;
- (2') B is better than A+, when the option set is  $S = \{A, A+, B\}$ ;
- (3') A+ is not worse than A, when the option set is  $S = \{A, A+, B\}$ .

This would spell big trouble, since (1'), (2'), and (3') would together constitute an instance of what I call "intransitivity within an option set" as opposed to the "intransitivity across different option sets" that we witness in propositions (1\*), (2\*), and (3\*). From the perspective of practical reason, intransitivity *within* an option set would be much the worse problem, since it could imply contradictory instructions and paralyze choice.

Fortunately, I argue, we are not forced to this conclusion, since in a three-possible case proposition (3') is false. The argument for this, in a nutshell, goes as follows: In A+, there is inequality between the A-people (the people who also exist in A) and the "new people", created through mere addition. Does this inequality make A+ worse than A? In *Reasons and Persons*, Parfit provides an argument (which he later repudiated, but which I still find plausible), according to which (roughly) if the only way of removing the inequality between the A-people and the new people is for the latter group never to exist, then the inequality is not a bad-making feature of A+. This is true in a two-possible choice between A and A+, and supports the judgment that, in *that* case, A+ is not worse than A. However, in a three-possible case, there *does* exist a way of removing the inequality between the A-people and the new people other than by preventing the new people from coming into existence, namely option B. B removes the inequality by making the new people better off, while reducing the wellbeing of the A-people by a lesser degree. Doing so would also increase

overall wellbeing. For this reason, there would be a complaint of avoidable *injustice* against A+ in the three-possible case that didn't apply in a two-possible choice between A and A+. And it is plausible that this injustice makes A+ a worse outcome than A in the three-possible case. (At least it will seem so to you if, like me, you accept the basic thought that the existence of more happy people, in itself, does not make the world better).

If this argument goes through, I have shown how we can make sense of each of our pairwise intuitions in the Mere Addition Paradox, without incurring the cost of intransitivity within an option set. In a three-possible case, we can affirm the following transitive ranking over the three options:  $A > B > A+$ . This solves the Mere Addition Paradox, and blocks the argument towards the Repugnant Conclusion.

A word on how my account of the Mere Addition Paradox differs from Temkin's seminal discussion in "Intransitivity and the Mere Addition Paradox" and *Rethinking the Good*: According to Temkin, what may underlie our intransitive pairwise intuitions is the fact that certain values, such equality, may be what he calls "essentially pairwise comparative". According to such a view of equality, "[t]here is no fact of the matter as to how bad the inequality in A+ *really* is considered just by itself. How bad it is depends upon the alternative compared to it. Compared to B, A+ is bad; compared to A, it isn't."<sup>3</sup> On this view, how A+ is to be judged in terms of equality is essentially *a function of the alternative it is being compared to*. If this were indeed the case, then how A+ compares to A in terms of inequality would *not* differ depending on whether or not B was also in the option set or not. Rather, in *both* a two-possible and a three-possible choice, A+ would not be worse than A in terms of equality, whereas it would be worse than B in the three-possible choice. In that case, however, our intransitive judgments in two-possible choices *would* carry over to a three-possible choice between A, A+, and B, thus raising the specter of intransitivity *within* an option set.

My account of the Mere Addition Paradox avoids this unattractive implication. The comparative goodness of A and A+ is *not*, as Temkin maintains, essentially pairwise comparative, but rather essentially *context-dependent*. Whether or not A+ contains injustice is a function of *what other outcomes are in the option set*. If only A and A+ are in the option set, there is no injustice in outcome A+. By contrast, given the presence of B in the option set, the inequality in A+ constitutes an injustice. This injustice matters both in comparisons with B *and with A*. That is, A+ is worse than B, on account of its injustice. *But it is also worse than A*, and for the same reason: A+ contains avoidable injustice, whereas A does not. Contra Temkin, what is challenged by the Mere Addition Paradox is *not* the transitivity of "all-things-considered better than" (in the sense that we care about, namely transitivity *within* an option set), but the Independence of Irrelevant Alternatives Principle of rational choice theory.

The fifth paper in the series, "**Future Persons and Victimless Wrongdoing**", published in German (with a response by Derek Parfit) as "Zukünftige Personen und Schuld ohne Opfer" in Markus Rüter and Sebastian Muders (eds.) *Worauf es ankommt: Derek Parfits praktische Philosophie in der Diskussion* (Hamburg: Felix Meiner Verlag, 2017) connects my work in population ethics with another central research interest of mine, namely the study of relational normativity (described in more detail in the next section of this document). Many moral

---

<sup>3</sup> Larry Temkin, "Intransitivity and the Mere Addition Paradox", *Philosophy & Public Affairs* 16.2 (1987), p. 150.

philosophers (e.g. Scanlon, Wallace) are attracted to a view of the core domain of *interpersonal* morality on which our moral requirements and prohibitions have a *relational* basis. On this picture, moral requirements have a constitutive connection with the claims, or entitlements, or possible complaints of another party. By the same token, for an act to be wrong, there must be someone with a legitimate claim that one not so act; this person would also be *wronged* if I violated their claim against me.

My paper argues that the Non-Identity Problem represents a counter-example to the view that an act can be wrong only if there is someone who is wronged by it. I examine various attempts to explain the wrong done in non-identity cases in terms of its effects on the person created, either in terms of non-comparative harms to the person or the violation of her rights. I argue that at least some instances of the non-identity problem, such as those involving the knowing creation of children with congenital conditions like deafness, don't admit of such a treatment. Instead they constitute a species of 'victimless wrongdoing', i.e. wrongdoing that doesn't involve anyone being wronged.

Finally, a sixth paper, co-authored with Adam Lerner, \***"Speciesism and Acceptable Lives"** (soon to be submitted to *Noûs*) examines some of the connections between my views on population ethics and the field of animal ethics. Specifically, we ask: should the same moral standards that govern the creation of human beings also apply to the creation of non-human animals, or are the relevant standards different? Consider: Just one in twenty rats born in the wild make it past their first birthday. The vast majority end up in the jaws of some predator or the clutches of some communicable disease within the first few months of life. Few of us think this is worth losing sleep over. But imagine circumstances in which only one in twenty human babies made it past their first birthday. (Suppose that procreation takes place in a refugee camp in which infant mortality is rampant). Such a state of affairs would keep us up at night. More than that, we are inclined to think that we have weighty moral reasons not to allow a human child to be born into circumstances where its life expectancy would only be measured in months. This would be so even if the infants had the same distributions of personal goods and bads (e.g., pleasures and pains) in their lives as the typical rat, and even if their lives contained more good than bad.

Why should we have such different attitudes towards rats and the unfortunate infants? It isn't obvious why our judgments of the two should differ when their lives seem so similar. This discussion brings out a tension between three claims, each of which find many adherents among contemporary philosophers, but which form an inconsistent triad:

- (1) **High Standards for Humans:** We have moral reason not to allow to come into existence a human child whose life contained the same distribution of personal goods and bads as the life of a typical rat, even if the goods outweighed the bads. That is, such a 'rat-like' life would not constitute an acceptable life (i.e., a life that we do not have moral reason not to allow to come into existence) for a human child. We should be anti-natalists with respect to humans with rat-like lives.
- (2) **Low Standards for Animals:** We do not have moral reason not to allow to come into existence a rat, or a similar non-human animal with a similar distribution of personal goods and bads, assuming the goods outweighed the bads. That is, such a life would constitute an

acceptable life for such an animal. We need not be anti-natalists with respect to non-human animals like rats.

- (3) **Anti-Speciesism about Acceptability Standards:** The standard for what constitutes an acceptable life is the same for humans and for non-human animals like rats.

These claims cannot all be true. Any two rule out the third. The choice is not obvious, because each claim enjoys support from philosophers and a good deal of *prima facie* plausibility.

After briefly considering the options of rejecting either High Standards for Humans or Low Standards for Animals, the main business of the paper is to sketch a novel account of why a life worth living might be unacceptable for a human but acceptable for an animal. Our account appeals to certain modal differences between humans and non-human animals. A human with a rat-like life has a life that has the modal property of being *tragic*, in the sense that in most nearby possible worlds in which she exists her life goes much better than it does in the actual world. Given the facts about species norms for longevity, death at one year is extremely premature for a human being. By contrast, the death of a rat in the wild at age one is not very premature, at least in absolute terms. Even excluding predators and communicable diseases, biological limitations on its lifespan mean that a rat typically will not live much beyond 21 months. We argue that, all else equal, we have *pro tanto* moral reason to avoid the creation of lives that will foreseeably be tragic, even if these lives are worth living. This is why creating a human being who will foreseeably die by age 1 is unacceptable, whereas there is no *pro tanto* moral reason to avoid creating a similarly short-lived rat.

We argue that allowing the assessment of the actual world to depend on its modal properties in this way is considerably more plausible than philosophers have realized, and that this is a significant result in value theory that goes beyond merely solving the puzzle at hand. If we are right, how we should think about the value of what actually happens is often powerfully influenced by what might have been. We distinguish this claim from superficially similar views already in the literature (e.g., Shelly Kagan's modal personism and Larry Temkin's essentially comparative view of moral ideals) and conclude by drawing out its potential to illuminate a number of other issues in normative ethics: the Mere Addition Paradox, the Non-Identity Problem, and the morality of exploitation.

### 3. "Interpersonal" Ethics

A third core research interest concerns *interpersonal* notions in ethics. Attracted as I am to a broadly contractualist account of morality, I am interested in the concept of *interpersonal justification* which is central to this theory. My article "**What We Owe to Hypocrites: Contractualism and the Speaker-Relativity of Justification**" (*Philosophy & Public Affairs*, 2016) explores the idea that the justificatory force of a moral argument may vary depending on *who* is addressing the argument to whom. An argument may serve to justify an action or policy when uttered by one speaker, while others may be unable to offer the same argument as a justification for the action in question. I argue that this "speaker-relativity" of moral justification has not received the attention from moral philosophers that it merits. Amongst other things, it raises novel problems for a number of prominent moral theories, most interestingly for Scanlonian

contractualism. Contractualists hold that an action is morally right if and because it is justifiable to each person (or, in the terms of Scanlon's formula, because it is 'licensed by a principle that no-one could reasonably reject'). Like attempts at justification, however, grounds for rejection must be presented in the form of arguments, addressed to the agent by those individuals who are adversely affected by his action. I argue that there are cases in which an action is intuitively wrong on account of its foreseeably bad consequences, yet *none* of those adversely affected have the ability to reasonably reject the principle that licenses its performance. To make this more concrete, consider one case from my paper:

*The Man Who Took Himself Hostage*: Jones, a single man with no dependents or friends, comes to your doorstep. Brandishing a pistol, he credibly threatens to kill himself unless you pay him \$20.

I maintain that failing to pay Jones under these circumstances is intuitively morally wrong, since by paying Jones, you have it in your power to prevent an awful moral tragedy at trivial cost to yourself. Yet Jones could not *reasonably reject* a principle that licenses you not to pay him; that is, Jones could not successfully justify to you that you morally ought to pay him in these circumstances. Jones has no legitimate basis for demanding that you pay him the ransom he demands; after all, even if you refuse to pay, it is entirely within Jones's power to avoid any adverse consequences for himself, by simply deciding not to shoot himself. This case, and others like it, threaten to drive a wedge between our moral intuitions and contractualism's account of moral rightness. They suggest that the supposed *analysans* (an action's being licensed by a principle that no-one could reasonably reject for personal reasons) and the *analysandum* (moral rightness) can sometimes come apart.

Besides investigating the phenomenon of speaker-relativity of interpersonal justification and leveraging this phenomenon to present a problem-case for Scanlonian contractualism, my article accomplishes two further things: First, I provide a detailed discussion of the nature of "interpersonal justification", which I distinguish from a less demanding notion of "propositional justification". (My paper is the first detailed analysis of the concept of interpersonal justification in the literature). Interpersonal moral justification, I maintain, is a form of *joint deliberation*. For it to be true that B successfully justifies to A that A ought to do  $\phi$  in circumstances  $c$ , this must be (part of) the correct answer to the question "what do *we*, A and B, have reason to do in circumstances  $c$ ?" That is, for B to successfully justify to A that he ought to do  $\phi$  in circumstances  $c$ , it must be the case that A cannot counter-propose a *different* course of action, jointly accessible to A and B, such that (i) this course of action does not involve A doing  $\phi$  and (ii) this course of action is a more plausible answer to the question "what are *we*, A and B, to do in circumstances  $c$ ?" In the *Man Who Took Himself Hostage* case above, Jones fails to provide you with a successful moral justification why you ought to pay him since, besides the course of action where you pay Jones the ransom and he does not kill himself, and the course of action where you do not pay Jones and he does kill himself, there is a third course of action jointly available to you and Jones, namely one where you do not pay Jones and yet he does not shoot himself. Moreover, this is a counter-proposal which Jones could not reasonably reject, since it secures the same good for Jones as the first course of action (the preservation of his life), but without imposing on you the burden of losing \$20.

Second, my article proposes a constructive revision to Scanlonian contractualism, which promises to solve well-known problems for the view, while differing from the revisions proposed by other sympathetic critics such as Derek Parfit in *On What Matters*. Both Parfit and I agree that, as stated in *What We Owe to Each Other*, Scanlon's contractualist formula does not constitute a fully adequate account of what it is for actions to be right or wrong in interpersonal morality. The formula yields intuitively incorrect verdicts in certain cases. These include situations where aggregative reasoning is intuitively appropriate (some of which I discuss in "Contractualism and Social Risk"), Parfit's Non-Identity Problem, and also cases like *The Man Who Took Himself Hostage* that I discuss in "What We Owe to Hypocrites". In response to these difficulties, Parfit pursues a strategy of revising contractualism "from the inside": we can retain the claim that contractualism provides an account of what it is for actions to be right or wrong in interpersonal morality, but only by making far-reaching revisions to the content of the view, such as jettisoning the so-called "Individualist Restriction", which prohibits interpersonal aggregation, or the "Personalist Restriction", which prevents us from appealing to purely *impersonal* considerations in rejecting moral principles. These revisions can "save" contractualism from the problem cases mentioned above, but at the price of stripping it of some of its most distinctive elements.

By contrast, I propose to revise contractualism "from the outside", by scaling back its ambition as a moral theory. On this proposal, Scanlon's formula remains as it is; however, instead of providing a complete account of what it *is* for actions to be right or wrong, the contractualist formula is now understood as capturing an important class of pro tanto moral reasons that contribute to *making* actions right or wrong, but that do not by themselves determine an action's rightness *all things considered*. This revision would allow Scanlon to maintain that there are other properties *besides* justifiability to each person – such as the goodness or badness of our action's consequences – which *independently* contribute to making our actions right or wrong. This "pluralist" proposal offers a straightforward solution to all the problem cases mentioned above. Moreover, unlike Parfit's proposal, it allows us to hold on to Scanlon's insight that there is a *distinctive* kind of objection to actions that fall afoul of his contractualist formula, as opposed to actions that are intuitively wrong for other reasons (e.g. because they produce impersonally worse consequences).

Another red thread that runs through my work on interpersonal ethics is trying to get clearer on the relationship between *monadic* moral notions, such as "is wrong" or "is blameworthy", which evaluate agents or their action, and *dyadic* or *interpersonal* moral notions that *link* an agent and a patient, such as "X *wrongs* Y".<sup>4</sup>

In my paper "**Dilemmas, Luck, and the Two Faces of Morality**" (under review at *Ethics*), I put this distinction to work in offering a new and unified solution to the problems of moral dilemmas and moral outcome luck. Proponents of moral dilemmas and proponents of moral outcome luck both implicitly endorse a thesis I call

**Parochial evaluation:** In evaluating a given action  $a_i$  or the agent who performed it, it is sometimes appropriate to "screen off" portions of modal space or at least to "downplay" their evaluative significance. In the case of moral outcome luck, this is the space of *alternative possible outcomes*  $\{o_1; o_2; \dots o_n\}$  that the action could have resulted in. In the case of moral dilemmas, this is the space of *alternative possible*

---

<sup>4</sup> See also my paper "Future Persons and Victimless Wrongdoing" described in Section 2.

actions  $\{a_1; a_2; \dots a_n\}$  that were available to the agent. Instead, we should make our assessment of the action or the agent exclusively or at least primarily a function of facts about  $a_i$  or  $o_i$  themselves.<sup>5</sup>

Proponents of moral outcome luck, I show, are parochialists about *outcomes*. They hold that an agent's degree of blameworthiness for performing a risky action  $a_i$  is not primarily a function of what possible outcomes  $\{o_1; o_2; \dots o_n\}$  the action *could* have resulted in. Rather, what matters crucially are facts about the outcome  $o_i$  which *actually* results. Proponents of moral dilemmas, typically subscribe to parochialism about *options*. Moral dilemmas arise when we decouple the question of whether a given course of action  $a_1$  is wrong from the wider context of choice – what alternative courses of action the agent could have chosen instead in that situation. Rather,  $a_1$  is wrong just because of what it is *in itself*. That the only alternative to  $a_1$  is  $a_2$  – which is *also* terrible – does not make it any less the case that it is wrong to do  $a_1$ . Apply the same reasoning to  $a_2$  and we arrive at a moral dilemma.

In the remainder of the paper, I then turn to two substantive ethical questions: should we accept that there are moral dilemmas, and should we accept that there is moral outcome luck? There are well-known and powerful objections to answering either question in the affirmative. Embracing moral outcome luck seems to fly in the face of the powerfully intuitive Control Principle, according to which an agent cannot be morally assessable for features of his action that were not under his control. And accepting that there are genuine moral dilemmas has struck many philosophers as extremely unpalatable, both on formal and on moral grounds.

Despite these misgivings, some philosophers have felt compelled to concede the possibility of genuine moral dilemmas and of moral outcome luck, for reasons that are, again, instructively parallel. Only by admitting moral dilemmas and moral outcome luck, they argue, can we make sense of certain aspects of our ordinary ethical experience – for instance, the very different *reactive attitudes* that seem appropriate in purported cases of moral outcome luck, depending on how things turn out, or the sense that in cases of tragic moral conflict, no matter how the agent acts, it will be appropriate to feel guilt for what he has done. I refer to these, collectively, as *arguments from phenomenology*, since they maintain that only by allowing the possibility of genuine moral dilemmas and genuine moral outcome luck can we ratify the characteristic moral emotions that seem apt in such cases.

Building on the notion of parochial evaluation, I present a novel strategy for solving the twin philosophical problems of moral dilemmas and moral outcome luck. The key that unlocks *both* problems, I argue, lies in clearly differentiating what I call “the two faces” of morality: We are both the *agents* and the *patients* of actions. We are active in the world, and are in turn affected by the activity of others. Accordingly, I argue, we must distinguish two perspectives from which we can morally evaluate a given action: an *agent-* and a *patient-centric* perspective. From the agent-centric perspective, the central question is “Has the agent comported herself well or poorly, morally speaking?” Facts about the rightness or wrongness of the agent's action belong to this first, agent-centric perspective. Also, on some views of blameworthiness, questions about the

---

<sup>5</sup> This is not, of course, to claim that any moral view that admits of moral dilemmas also admits of moral outcome luck, or vice versa. It is only to claim that moral theories that give rise to either phenomenon have a structural commonality, namely a commitment to parochial evaluation.

agent's degree of blameworthiness are exclusively the province of this agent-centric perspective of evaluation. By contrast, the central question from the patient-centric perspective is "How have people been *affected* by the agent's action? Specifically, has anyone been *wronged* by the agent's behavior? If yes, how seriously?"

The key claim of the second half of the paper is this: Bracketing the question whether we should accept parochial evaluation with regard to *agent-centric* questions about moral rightness or blameworthiness, it is extremely plausible that when it comes to the *patient-centric* question whether someone has been *wronged* by the agent's action (and if so, how seriously), we *should* be parochialists, both (i) about outcomes and (ii) about options. I show that how seriously a person has been wronged by an action depends to a large degree, not on what outcome the action *could* have produced, but on what outcome *actually* occurred. Moreover, I argue, there are tokens of certain extreme act-types (killing, torture, rape) which can be said to wrong a patient, *no matter* what alternatives the agent had to treating her in this way.

However, once we realize that we have to be parochialists about *wronging*, we can use this insight to solve the problem of moral dilemmas and defang the problem of moral luck. I argue that the distinctive phenomenology of tragic moral conflicts can be fully accounted for in terms of parochialism about *wronging*. There is no need to postulate genuine moral dilemmas in order to explain why, in cases of tragic moral conflict, an agent can rightly believe that, no matter how he acts, he will have reason to feel guilty. With regard to the problem of moral outcome luck, I show that what accounts for the different reactive attitudes that seem appropriate vis-à-vis lucky and unlucky agents are not *agent-centric* facts about how well or poorly the agent has comported herself, but *patient-centric* facts about how seriously other people have been *wronged* by the agent. Depending on the account of blame we endorse, this implies that either there is no moral luck or, if there is, it is not of a kind that conflicts with the Control Principle.

Alongside "Contractualism and Social Risk" and "Conditional Reasons and the Procreation Asymmetry", I consider this my most important paper to date. The distinction between agent- and patient-centric evaluation allows us to make sense of the thought that our ethical lives are shot through with contingency and tragedy, without having to gainsay our ability to act well even under the most difficult of circumstances. Morality, I argue, is Janus-faced: In our capacity as moral decision-makers, striving to act rightly and to avoid being at fault, we are in control of our destinies: I claim that there is no need to accept the paradoxical thought that life and morality can conspire to place us in situations where we cannot but act *wrongly*. And even when we do comport ourselves poorly, by taking morally unjustifiable risks, how badly this impacts our records as moral *agents* is a function only of factors that were under our control.

Yet, in our capacity as inhabitants of a social world, who can relate to one another in various roles, including that of *wrongdoer* and *victim*, there is scope for great contingency and tragedy. When it comes to how seriously we have wronged another person by taking an unjustified risk, we are often utterly at the mercy of fate. And although there may be no genuine moral dilemmas, I contend that life can nonetheless place us in situations from which we cannot escape without, in one way or another, feeling a profound sense of guilt. As I put it in the paper, we can have tragedy without the price of paradox.



#### 4. National Partiality and the Ethics of Immigration

My views on the nature of interpersonal justification have important implications for topics in political philosophy. I begin to explore some of these in my paper “**National Partiality, Immigration, and the Problem of Double-Jeopardy**” (forthcoming in *Oxford Studies in Political Philosophy*, Volume 6).

The foundational conviction of contemporary liberal thought is that all persons possess equal worth and are entitled to equal concern and respect by others. At the same time, nation states, as the primary organs of our collective self-governance, frequently pursue policies that are strikingly *partial* towards the interests of compatriots over those of foreigners. A common strategy for justifying this national partiality is to view it as grounded in *associative* obligations that we incur by standing in special relationships with our fellow citizens. For instance, according to familiar arguments from ‘fair play’, voluntarily partaking in and benefiting from cooperative relationships with our fellow citizens gives us reciprocal obligations to maintain these social institutions and care for needy compatriots. By contrast, we do not have corresponding associative obligations towards foreigners. I maintain that such arguments from associative obligations face an important stumbling block. I argue for the following “Boundary Principle”, according to which we cannot appeal to special relationships amongst members of a group to justify an attitude of partiality towards them unless the boundaries of this group, and the fact that they *exclude* other persons from membership, can also be justified. If this is right, arguments from associative obligations, by themselves, are necessarily incomplete. Most modern states severely limit the scope of economic migration, and thereby prevent many would-be contributors to the collective enterprises of affluent societies from *entering into* those special relationships with us that would give rise to associative obligations on our part. Not only are would-be immigrants excluded from the benefits that participation in the cooperative schemes of affluent nation-states would *itself* confer; in addition, the fact that they do not stand in such cooperative relationships with us is then appealed to in order to justify the fact that we assign significantly less weight to their interests in the formulation of state policy. The upshot of my discussion is not that a defense of national partiality in terms of associative obligations cannot succeed. It is rather that a successful defense of national partiality in these terms is more closely tied to the question whether principled restrictions can be placed on immigration than most political philosophers have hitherto acknowledged.

#### 5. Self-Knowledge and Rationality

Finally, I also have a strong interest in questions of practical rationality, and requirements of *structural* rationality in particular. In my manuscript “**Self-Knowledge and Structural Rationality**”, I argue that making sense of certain requirements of structural rationality (in particular, the rational requirement not to believe propositions that are logically inconsistent and the rational requirement to intend the necessary means to my ends) requires us to adopt a particular epistemology of the first person. It is striking that, though the requirements of structural rationality govern combinations of *attitudes*, there has been little discussion of how questions regarding the authority of structural rationality might relate to the question of how we know our own minds. I

argue that this is a mistake. Far from being neutral in their epistemological commitments, requirements of structural rationality have their home in a particular epistemology of the first person. Making sense of requirements of structural rationality (both of their *felt* authority from the first-personal perspective and of the actual normative status of these requirements), requires us to adopt something like the ‘deliberative’ model of self-knowledge, defended most prominently in the work of Dick Moran. According to this view, the canonical way in which I know my beliefs and intentions is not through introspection or observational evidence, but by *making up my mind* about what to believe and to intend.

## 6. What is my best work to date?

If I myself had to rank the twelve pieces that I am submitting as part of this tenure file in rough order of quality, the ranking would look like this:

1. “Contractualism and Social Risk”
2. “Dilemmas, Luck, and the Two Faces of Morality”
3. “Conditional Reasons and the Procreation Asymmetry”
4. “What We Owe to Hypocrites”
5. “National Partiality, Immigration, and the Problem of Double-Jeopardy”
6. “On the Survival of Humanity”
7. “Chancy Causation and the Problem of Aggregate Events”
8. “Zukünftige Personen und Schuld ohne Opfer” (“Future Persons and Victimless Wrongdoing”)
9. “Treatment versus Prevention”
10. “Uncertainty and Justifiability to Each Person”
11. “Context-Dependent Betterness and the Mere Addition Paradox”
12. “Prioritarisme” (handbook entry co-authored with Ekédi Mpondo-Dika)

## 7. Professional recognition and impact of my work

I am fortunate in the recognition that my research has received to date. My doctoral dissertation *‘Making People Happy, Not Making Happy People’: A Defense of the Asymmetry Intuition in Population Ethics* was awarded Harvard University’s Emily and Charles Carrier Prize “for an outstanding doctoral thesis on a subject in social, political, or moral philosophy”. My article “Contractualism and Social Risk” won the American Philosophical Association’s Gregory Kavka Prize in 2017, for the best paper in political philosophy published in the two preceding years.

There have been two workshops wholly devoted to my research. In the fall of 2017, the Einstein Ethics Group in Berlin, under the aegis of Professor R.J. Wallace (Berkeley), organized a [two day workshop](#) around my work in Berlin. Commentators at that workshop included Christoph Fehige (Saarbrücken), Tim Henning (Stuttgart), Ulrike Heuer (Leeds), Erasmus Mayr (Erlangen), Kirsten Meyer (HU Berlin), Lukas Meyer (Graz), Véronique Munoz-Dardé (UCL/Berkeley), Juri Viehoff (Zürich), and Tatjana Višak (Mannheim). In the spring of 2019, the Edmond J. Safra Center for Ethics at Harvard University, at the initiative of Danielle Allen, held a [one-day symposium](#) on

some of my papers in progress. Commentators on that occasion included Caspar Hare (MIT), John John Goldberg, Rebecca Henderson, Priyanka Menon, Gina Schouten, and Lucas Stanczyk (all Harvard).

In addition, I have been invited to give lectures at many leading centers of philosophical excellence, including Berkeley, Columbia, Duke, Harvard, the Hebrew University of Jerusalem, the Institute for Futures Studies in Stockholm, the Institute of Philosophy in London, MIT, Oxford University, NYU, Rutgers, Stanford, UNC-Chapel Hill, UPenn, USC, the University of San Diego, UT Austin, the University of Toronto, and the University of Warwick. I was also part of a group of 20 philosophers and economists who consulted on ethical issues relating to the 2010 Global Burden of Disease Report, the most comprehensive survey of the world's health.

I have also been asked to serve on the board of recognized publications: I am an advisory board member of the *Zeitschrift für Ethik und Moralphilosophie*, the leading German-language journal for value theory, and recently became an Associate Editor at the *Journal of Moral Philosophy*.

Finally, in 2018 I was appointed to a Bicentennial Preceptorship at Princeton, a special distinction which is bestowed on one to two assistant professors from all the humanities and social sciences each year.

Turning to the impact that my work has had so far: The article of mine that has had the largest influence to date is “Contractualism and Social Risk”. It has been cited over fifty times and helped spark a lively literature on questions of contractualism, risk, and aggregation, with responses to my article appearing in *Philosophy & Public Affairs*, *Ethics*, *The Journal of Moral Philosophy*, *Ethical Theory & Moral Practice*, *Utilitas*, *The Journal of Ethics and Social Philosophy*, and *The Journal of Legal Analysis*, as well as in R.J. Wallace's book *The Moral Nexus* (PUP, 2019). In connection with the awarding of the Gregory Kavka Prize 2017, for the best paper in political philosophy published in the two preceding years, there was a symposium devoted to my article at the 2017 Pacific APA conference, at which Rahul Kumar and Michael Blake gave comments and I responded. T.M. Scanlon, the originator of moral contractualism, remarks in a recent article (“Reply to Zofia Stemplowska,” *Journal of Moral Philosophy* 10 (2013)), that his earlier opposition to ex ante contractualism was “a mistake,” and credits my article for changing his mind.

“Contractualism and Social Risk” has also attracted attention from senior legal scholars, including Barbara Fried, Gregory Keating, John Oberdiek, and Dov Waisman. It was also one of four works discussed at a workshop at the University of San Diego School of Law's Institute for Law & Philosophy on “Individual v. Statistical Lives, Ex Ante v. Ex Post,” organized by Larry Alexander in 2017. The other panelists were Richard Arneson (UC San Diego), Caspar Hare (MIT) and Matthew Adler (Duke Law).

“Uncertainty and Justifiability to Each Person”, although it has a more specialist audience, has been cited and discussed in pieces in *Ethics*, *Analysis*, *The Journal of Applied Philosophy*, *The Journal of Practical Ethics*, and *The Oxford Handbook of Distributive Justice*, as well as in Shlomi Segall's book *Why Inequality Matters* (CUP, 2016).

“Treatment versus Prevention in the Fight Against HIV/AIDS and the Problem of Identified versus Statistical Lives” has had an impact both on the philosophical literature, where it has been discussed by ethicists concerned with the normative relevance of the ‘identified vs statistical lives’ distinction (papers in *Utilitas*, *The Journal of Ethics and Social Philosophy*) but also on the literature on public health ethics and even on the clinical literature (*Journal of Clinical Ethics*, *Public Health Ethics*, *Journal of Acquired Immune Deficiency Syndromes*). As the first philosopher to provide a thorough ethical examination of the “Treatment as Prevention” paradigm, I was also invited to a number of Harvard Medical School conferences in 2013, where I discussed my work with AIDS researchers in the field, including Max Essex (Harvard Medical School) and Till Bärnighausen (Heidelberg), who at the time were conducting epidemiological field trials in KwaZulu-Natal on the effectiveness of treatment-as-prevention. It was also a focal point of discussion at an APA panel on “Justice and Risk-Distribution” at the 2015 Pacific APA, which I organized under the aegis of the APA’s Committee on Philosophy and Medicine. The speakers at the panel were Paul Kelleher, Rahul Kumar, Colleen Murphy, and John Oberdiek. It was also discussed, together with “Contractualism and Social Risk”, at the workshop on “Individual v. Statistical Lives, Ex Ante v. Ex Post” organized by Larry Alexander at the University of San Diego’s Institute for Law & Philosophy.

Although “Chancy Causation and the Problem of Aggregate Effects” still has an R&R, a precursor of this paper (“Causal Dependence and Chance: The New Problem of False Negatives”), which focused on the Problem of Aggregate Effects, already received a two-page discussion in Ned Hall and Laurie Paul’s *Causation: A User’s Guide*, the most authoritative recent handbook on philosophical theories of causation. Hall and Paul write that it is “a lovely paper,” which raises one of the more serious problems for Lewis’s account of probabilistic causation.

My work in population ethics is well known to most of the leading experts in this field and has been favorably received. In his published response to “Zukünftige Personen und Schuld ohne Opfer” (Future Persons and Victimless Wrongdoing), the late Derek Parfit calls it an “outstanding paper” and writes that he “accepts all of Frick’s main claims”. Jeff McMahan, one of the leading contemporary experts in population ethics, wrote to me about “Conditional Reasons and the Procreation Asymmetry”: “your paper on the Asymmetry has influenced my thinking more than any other work in population ethics since *Reasons and Persons*. It has prompted and guided me in going deeper into these issues than I would have been able to do otherwise. I am still deeply uncertain what to believe but thinking long and hard about the arguments in your paper has helped me more than anything else” (personal communication, Dec. 2019). R. J. Wallace also discusses both these papers in *The Moral Nexus*. My paper “On the Survival of Humanity” is cited approvingly by Sam Scheffler in *Why Worry About Future Generations?* and is also discussed in a number of recent papers on existential risk.

In addition, I have been an invited speaker at specialist conferences on population ethics at the Centre for Futures Studies in Stockholm in 2016 (organized by Gustaf Arrhenius), at a panel at the 2016 Eastern APA (organized by Melinda Roberts), and at a conference at Duke University (organized by Matt Adler) in 2017. Finally, in May 2019, Tina Rulli (UC Davis) and I co-organized a large international workshop on “[New Work on Population and Procreation Ethics](#)“, which brought together some of the leading international figures in this area. Last but not least, of the

four Fulbright Scholars whom I have hosted at Princeton, three have come explicitly in order to work with me on population ethics.

My work on the speaker-relativity of justification has also been cited and discussed by leading moral philosophers, such as Roger Crisp, Tom Dougherty, Kasper Lippert-Rasmussen, Brendan de Kenessey, T.M. Scanlon, and Alec Walen. And, though not yet published, my work on moral dilemmas, in addition to my published papers on risk, have brought me into conversation with some of the leading scholars in Just War Theory. For instance, Victor Tadros asked me to co-host a workshop (together with Helen Frowe, University of Stockholm) on his forthcoming book *To Do, to Die, to Reason Why* at the UCHV this fall, at which I will also be one of five speakers.

I am especially proud that my work in philosophy has attracted interest beyond the boundaries of our discipline. For instance, in October 2016, I was a speaker at an important interdisciplinary conference on Normative Ethics and Welfare Economics, which brought together established senior scholars from both disciplines (among the economists, the Nobel laureate Al Roth, Greg Mankiw, and Marc Fleurbaey; among the philosophers, Michael Sandel, Dan Hausman, Caspar Hare, and Tim Scanlon) and a few up-and-coming junior scholars. I myself was paired up on a panel on intergenerational ethics and welfare economics with Partha Dasgupta, the Frank Ramsey Professor Emeritus of Economics at the University of Cambridge, and our commentator was Glen Weyl, Principal Researcher at Microsoft Research.

Likewise, I am starting to build a network of interlocutors in the world of legal academia. Besides the aforementioned workshop on “Individual v. Statistical Lives, Ex Ante v. Ex Post” at the University of San Diego School of Law in 2017, and published discussions of my work on risk by legal scholars such as Barbara Fried, Gregory Keating, or John Oberdiek, I have also spoken at Columbia Law School (at the invitation of Joseph Raz), at Cornell Law School (at the invitation of Andrei Marmor and Emad Atiq), and at the Law and Philosophy Workshop at UPenn (at the invitation of Mitchell Berman and Stephen Perry). I am also a member of a criminal law and philosophy discussion group, run by Michael Moore and Heidi Hurd (University of Illinois), which counts amongst its members Kimberly Ferzan, Doug Husak, Leo Katz, John Oberdiek, Jeremy Waldron, Alec Walen, and Gideon Yaffe. The group meets one to two times a year for symposia at the intersection of philosophy and law (our next meeting is this May in Surrey, UK, for a workshop on threshold deontology).

Finally, I am actively exploring the relevance of some of my recent work on dyadic vs monadic normativity, and on causal and moral responsibility for aggregate effects, for the law. In the wake of the workshop on my research at the Harvard Safra Center, I had a long correspondence with John Golberg, Carter Professor of General Jurisprudence at Harvard Law School, on how the distinction between “agent-“ and “patient-centric” modes of evaluation that I draw in “Dilemmas, Luck, and the Two Faces of Morality”, maps onto the distinction between criminal and tort law, and how my arguments can help us explain many jurists’ sense that outcome luck, while problematic with regard to *criminal* punishment, is par for the course in the domain of tort law. Last but not least, just a few days ago, I was contacted by Alexandra D. Lahav, Ellen Ash Peters Professor at the University of Connecticut School of Law, who had read my paper on “Chancy Causation and the Problem of Aggregate Events” in researching a project on probabilistic causation in tort law, especially in mass harm situations, and is eager to discuss more of my work

in progress on these topics. She writes that “it is very generative for what I am trying to do and I haven’t found anything else like it.”

## Appendix: All Citations of my Work, By Area of Research and By Paper

My papers appear in bold. A star (\*) before a paper title denotes that the paper contains a discussion of my work. Papers without this symbol merely cite my work.

### I. ETHICS OF RISK

**“Contractualism and Social Risk”, *Philosophy & Public Affairs* 43.3 (2015): 175-223.**

Published and forthcoming:

Ashford, Elizabeth and Mulgan, Tim, “Contractualism”, *Stanford Encyclopedia of Philosophy*: <https://171.67.193.20/entries/contractualism/>

Byskov, Morten Fibieger. “Utilitarianism and Risk.” *Journal of Risk Research* (2018): 1-12.  
Christiansen, Andreas. “On the Cognitive Argument for Cost-Benefit Analysis.” *Ethical Theory and Moral Practice* 21, no. 2 (2018): 217-30.

Christiansen, Andreas. “On the Cognitive Argument for Cost-Benefit Analysis.” *Ethical Theory and Moral Practice* 21, no. 2 (April 1, 2018): 217–30. <https://doi.org/10.1007/s10677-018-9883-9>.

Eggert, Linda. “Harming the Beneficiaries of Humanitarian Intervention.” *Ethical Theory and Moral Practice* 21, no. 5 (2018): 1035-50.

\*Fried, Barbara H. “Facing up to Risk.” *Journal of Legal Analysis* 10 (2018).

\*Gordon-Solmon, Kerah. “Should Contractualists Decompose?” *Philosophy & Public Affairs* 47, no. 3 (2019): 259–87. <https://doi.org/10.1111/papa.12146>.

\*Holm, Sune. “The Luckless and the Doomed. Contractualism on Justified Risk-Imposition.” *Ethical Theory and Moral Practice* 21, no. 2 (2018): 231-44.

\*Horton, Joe. “Aggregation, Complaints, and Risk”, *Philosophy & Public Affairs* 45.1 (2017): 54-81.

Horton, Joe. “Aggregation, Risk, and Reductio.” *Ethics* (forthcoming).

James, Aaron, “Contractualism’s (Not So) Slippery Slope”, *Legal Theory*, Vol. 18 (2012), 263–292.

\*John, Tyler M., and Joseph Millum. “First Come, First Served?”, *Ethics* 130 (January 2020): 179-207.

Katz, Corey. "Contractualism, Person-Affecting Wrongness and the Non-Identity Problem." *Ethical Theory and Moral Practice* 21, no. 1 (2018): 103-19.

Keating, Gregory C. "Principles of Risk Imposition and the Priority of Avoiding Harm." *Revus. Journal for Constitutional Theory and Philosophy of Law/Revija za ustavno teorijo in filozofijo prava*, no. in print (2018).

Kirkpatrick, James R. "Permissibility and the Aggregation of Risks." *Utilitas* 30, no. 1 (2018): 107-19.

Kumar, Rahul. "Risking Future Generations." *Ethical Theory and Moral Practice* 21, no. 2 (2018): 245-57.

\*Kuyumcuoglu, Huseyin S. "Sweatshops, Harm, and Interference: A Contractualist Approach." *Journal of Business Ethics*, October 19, 2019. <https://doi.org/10.1007/s10551-019-04302-9>.

Lazar, Seth. "Anton's Game: Deontological Decision Theory for an Iterated Decision Problem." *Utilitas* 29, no. 1 (2017): 88-109.

\* \_\_\_\_\_ "Limited Aggregation and Risk." *Philosophy & Public Affairs* 46, no. 2 (2018): 117-59.

Loi, Michele, and Markus Christen. "Ethical Frameworks for Cybersecurity." In *The Ethics of Cybersecurity*, edited by Markus Christen, Bert Gordijn, and Michele Loi, 73–95. The International Library of Ethics, Law and Technology 21. Springer International Publishing, 2020. <https://doi.org/10.1007/978-3-030-29053-5>.

\*Mardellat, Victor. "Contractualism and the Paradox of Deontology." *Philosophical Studies*, January 7, 2020. <https://doi.org/10.1007/s11098-019-01406-w>.

Moehler, Michael. *Minimal Morality: A Multilevel Social Contract Theory*. Oxford University Press, 2018.

Mogensen, Andreas. "The Callousness Objection." In *Effective Altruism: Philosophical Issues*, edited by Hilary Greaves and Theron Pummer, 227-244. Oxford University Press, 2019.

\*Nissan-Rozen, Ittay. "How to Be an Ex-Post Egalitarian and an Ex-Ante Paretian." *Analysis* 77, no. 3 (2017): 550-58.

Oberdiek, John. *Imposing Risk: A Normative Framework*. Oxford University Press, 2017.

Oksanen, Markku, Ashley Dodsworth, and Selina O'Doherty. *Environmental Human Rights: A Political Theory Perspective*. Routledge, 2017.

Parr, Tom, and Adam Slavny. "What's Wrong with Risk?". *Thought: A Journal of Philosophy* (2019).



Rowe, Thomas. “Risk and the Unfairness of Some Being Better Off at the Expense of Others.” *Journal of Ethics and Social Philosophy* 16, no. 1 (2019).

Rowe, Thomas, and Alex Voorhoeve. “Egalitarianism under Severe Uncertainty.” *Philosophy & Public Affairs* 46, no. 3 (2018): 239-68.

\*Rüger, Korbinian. “On Ex Ante Contractualism.” *J. Ethics & Soc. Phil.* 13 (2018): 240.

\*Scanlon, T.M., “Reply to Zofia Stemplowska”, *Journal of Moral Philosophy* 10 (2013) 508–514.

\*Setiya, Kieran, “Other People.” Forthcoming in *The Value of Humanity*, edited by Sarah Buss and Nandi Theunissen, Oxford University Press.

Song, Fei. “Rights against High-Level Risk Impositions.” *Ethical Theory and Moral Practice* (2019): 1-16.

St. Clair, Robert. *Poetry, Politics, and the Body in Rimbaud: Lyrical Material*. Oxford University Press, 2018.

\*Suikkanen, Jussi. “Ex Ante and Ex Post Contractualism: A Synthesis.” *The Journal of Ethics* 23, no. 1 (2019): 77-98.

\*Waisman, Dov, “Equity and Feasibility Regulation”, *University of Richmond Law Review*, Volume 50 (May 2016), pp. 1263-1331.

\*Walen, Alec, “Proof Beyond a Reasonable Doubt: A Balanced Retributive Account”, *Louisiana Law Review*, Vol. 74 (2015), pp. 355-446.

\*Walen, Alec, “Risks and Weak Aggregation: Why Different Models of Risk Suit Different Types Cases,” forthcoming in *Ethics*.

\*Wallace, R.J., *The Moral Nexus* (Princeton: Princeton University Press, 2019).

\*Wilson, James. “VII—Internal and External Validity in Thought Experiments.” *Proceedings of the Aristotelian Society* 116, no. 2 (July 1, 2016): 127–52. <https://doi.org/10.1093/arisoc/aow008>.

Załoski, Wojciech. “The Moral Status of Helping and the Identified Victim Effect.” *Principia*, 2018, 49–68. <https://doi.org/10.4467/20843887PI.18.002.9885>.

\*Żuradzki, Tomasz. “The normative implications of the preference toward identified people”, *Diametros* 51 (2017): 113–136 (in Polish).

\*———. “The Normative Significance of Identifiability.” *Ethics and Information Technology* 21, no. 4 (December 1, 2019): 295–305. <https://doi.org/10.1007/s10676-018-9487-z>.

### **Dissertations and Working Papers:**

Algander, Per, “Contractualism and the Repugnant Conclusion.” Manuscript, Umeå University.  
[https://www.academia.edu/40186903/Contractualism\\_and\\_the\\_Repugnant\\_Conclusion](https://www.academia.edu/40186903/Contractualism_and_the_Repugnant_Conclusion)

Bijloos, Annemarie Dorothee. *Ethics and Risk. Toward a Responsible Approach to Acceptable Risk Impositions*. D.Phil Thesis, Stellenbosch University, 2017.

Bolinger, Renée Jorgensen. *Mistaken Defense and Normative Conventions*. Ph.D. Dissertation, University of Southern California, 2017.

Fonseca, Max. *Your Treatment, My Treat? On Lifestyle-Related Ill-Health and Reasonable Responsibility*. Ph.D. dissertation, Department of Political Science, Stockholm University, 2019.

Greaves, Hilary, and William MacAskill. “The Case for Strong Longtermism.” Manuscript, 2019. [https://globalprioritiesinstitute.org/wp-content/uploads/2019/Greaves\\_MacAskill\\_The\\_Case\\_for\\_Strong\\_Longtermism.pdf](https://globalprioritiesinstitute.org/wp-content/uploads/2019/Greaves_MacAskill_The_Case_for_Strong_Longtermism.pdf)

Healey, Richard, *The Power of Consent*. DPhil Thesis, University of Sheffield, 2016.

\*Scanlon, T.M., “Contractualism and Justification.” Manuscript, 2017. (To appear in the proceedings of the 2016 Lauener Conference.)  
[https://www.academia.edu/35861169/Contractualism\\_and\\_Justification.pdf](https://www.academia.edu/35861169/Contractualism_and_Justification.pdf)

Suikkanen, Jussi. “Ethical Theories as Methods of Ethics,” Arizona Normative Ethics Workshop, 2020.

Zuehl, Jake, *Collective Self-Determination*, PhD Dissertation, Princeton University (May 2016).

**“Uncertainty and Justifiability to Each Person: Response to Fleurbaey and Voorhoeve”, in Nir Eyal, Samia Hurst, Ole Norheim and Dan Wikler (eds.), *Inequalities in Health: Concepts, Measures, and Ethics* (New York: Oxford University Press, 2013).**

### **Published:**

Bärnighausen, Till, Nir Eyal, and Daniel Wikler. “HIV Treatment-as-Prevention: Research at a Crossroads.” *PLOS Medicine* 11.6 ((2014): e1001654.  
<https://doi.org/10.1371/journal.pmed.1001654>

Bognar, Greg. "Review of Nir Eyal, Samia Hurst, Ole Norheim and Dan Wikler (eds.), *Inequalities in Health: Concepts, Measures, and Ethics*, Oxford University Press, 2013." *Ethical Perspectives* 21.2 (2014): 312-315.

Kamm, Frances. "Cost Effectiveness Analysis and Fairness." *Journal of Practical Ethics* 3, no. 1 (2015).

Nebel, Jacob M. "An Intrapersonal Addition Paradox." *Ethics* 129, no. 2 (January 2019): 309–43. <https://doi.org/10.1086/700085>.

Nissan-Rozen, Ittay. "How to Be an Ex-Post Egalitarian and an Ex-Ante Paretian." *Analysis* 77, no. 3 (2017): 550-58.

Otsuka, Michael, and Alex Voorhoeve. "Equality Versus Priority." *The Oxford Handbook of Distributive Justice* (2018): 65.

\*Rowe, Thomas. "Risk and the Unfairness of Some Being Better Off at the Expense of Others." *Journal of Ethics and Social Philosophy* 16, no. 1 (2019).

Segall, Shlomi. *Why Inequality Matters*. Cambridge University Press, 2016.

\*Thompson, Christopher. "Rose's Prevention Paradox." *Journal of Applied Philosophy* 35, no. 2 (2018): 242-56.

\*Voorhoeve, Alex, and Marc Fleurbaey. "Priority or Equality for Possible People?." *Ethics* 126, no. 4 (2016): 929-54.

\*Wiesner, Valentin. "Risk Aversion Is Grounds for Rejecting the Ex Ante Pareto Principle." *Studies in Philosophy, Politics and Economics* 1, no. 1 (July 19, 2019): 34–38.

#### Working papers

\*Scanlon, T.M., "Contractualism and Justification." Manuscript, 2017. (To appear in the proceedings of the 2016 Lauener Conference.)

[https://www.academia.edu/35861169/Contractualism\\_and\\_Justification.pdf](https://www.academia.edu/35861169/Contractualism_and_Justification.pdf)

**"Treatment versus Prevention in the Fight Against HIV/AIDS and the Problem of Identified versus Statistical Lives" in Glenn Cohen, Norman Daniels, and Nir Eyal (eds.), *Identified versus Statistical Lives: An Interdisciplinary Perspective* (New York: Oxford University Press, 2015).**

#### Published

Eyal, Nir. “Pediatric Heart Surgery in Ghana: Three Ethical Questions.” *J Clin Ethics* 25, no. 4 (2014): 317-22.

\*Kelleher, J. Paul. “Relevance and Non-Consequentialist Aggregation.” *Utilitas* 26, no. 4 (2014): 385-408.

Mogensen, Andreas. “The Callousness Objection.” In *Effective Altruism: Philosophical Issues*, edited by Hilary Greaves and Theron Pummer, 227-244. Oxford University Press, 2019.

Ng, Jared, Manne Sjöstrand, and Nir Eyal. “Adding Lithium to Drinking Water for Suicide Prevention—the Ethics.” *Public Health Ethics* (2019).

Rüger, Korbinian. “On Ex Ante Contractualism.” *J. Ethics & Soc. Phil.* 13 (2018): 240.

Zhang, Yi, Till Bärnighausen, and Nir Eyal. “When Global Art Budgets Cannot Cover All Patients, Who Should Be Eligible?”. *JAIDS Journal of Acquired Immune Deficiency Syndromes* 81, no. 2 (2019): 134-37.

\*Żuradzki, Tomasz. “The normative implications of the preference toward identified people”, *Diametros* 51 (2017): 113–136 (in Polish).

\*\_\_\_\_\_. “The Normative Significance of Identifiability.” *Ethics and Information Technology* (2018): 1-11.

### Theses and Working Papers

Flynn, Spencer. *The Ethics of the Rule of Rescue: Guidelines for Use in the Medical Setting*. Honors Thesis, Department of Philosophy, Duke University.

## **“Chancy Causation and the Problem of Aggregate Effects” (R&R at *Philosophical Studies*)**

### Published

\*Hall, Ned, and L.A. Paul. *Causation: A User’s Guide*. Oxford University Press, 2013.

### Working Paper

Walen, Alec. “Probabilistic Partial Causation and a Gap in Criminal Law.”  
[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3231949](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3231949)

## **II. POPULATION ETHICS**

**“On the Survival of Humanity”, *Canadian Journal of Philosophy*, 47.2-3 (2017): 344-367.**

### Published

Ashford, Elizabeth, and Tim Mulgan. “Contractualism”, *Stanford Encyclopedia of Philosophy*.

Baum, Seth D, Stuart Armstrong, Timoteus Ekenstedt, Olle Häggström, Robin Hanson, Karin Kuhlemann, Matthijs M Maas, *et al.* “Long-Term Trajectories of Human Civilization.” *foresight* 21, no. 1 (2019): 53-83.

Gottlieb, Joseph. “Space Colonization and Existential Risk.” *Journal of the American Philosophical Association* (2019): 1-15.

Hopster, Jeroen. “The Speciesism Debate: Intuition, Method, and Empirical Advances.” *Animals* 9, no. 12 (2019): 1054.

Kaczmarek, Patrick, and Simon Beard. “Human Extinction and Our Obligations to the Past.” *Utilitas* (2020): 1-10.

\*Qizilbash, Mozaffar. “On Parity and the Intuition of Neutrality.” *Economics & Philosophy* 34, no. 1 (2018): 87-108.

Scheffler, Samuel. *Why Worry About Future Generations?* : Oxford University Press, 2018.

Smith, Holly M, and David E Black. “The Morality of Creating and Eliminating Duties.” *Philosophical Studies* (2019): 1-30.

Torres, Phil. “Facing Disaster: The Great Challenges Framework.” *foresight* 21, no. 1 (2019): 4-34.

### Theses and Working Papers

Carlsmith, Joseph. “Helping People, Creating People, and Shaping the Far Future.” Manuscript, New York University, 2017.

Kolodny, Niko. “Saving Posterity from a Worse Fate.” Manuscript (to be published in a *Festschrift* for Derek Parfit), 2020. <https://www.ocf.berkeley.edu/~ngkolodny/SPWF1.pdf>

Plant, Michael. *Doing Good Badly? Philosophical Issues Related to Effective Altruism*. D.Phil. Thesis, University of Oxford, 2019.

Southan, Rhys. “Life Has Value, but Non-Existence Does Not.” B.Phil. Thesis, University of Oxford, 2018.

Von Sass, Hartmut. “The Curious Case of Normative Incomparability: Comparisons, Animals, and the Quest for Adequacy.” Manuscript.

[https://www.academia.edu/37046184/The\\_Curious\\_Case\\_of\\_Normative\\_Incomparability\\_Comparisons\\_Animals\\_and\\_the\\_Quest\\_for\\_Adequacy](https://www.academia.edu/37046184/The_Curious_Case_of_Normative_Incomparability_Comparisons_Animals_and_the_Quest_for_Adequacy)

***‘Making People Happy, Not Making Happy People’: A Defense of the Asymmetry Intuition in Population Ethics.* Dissertation defended at Harvard University, September 2014.**

### Published

\*Earl, Jake. “A Portable Defense of the Procreation Asymmetry.” *Canadian Journal of Philosophy* 47, no. 2-3 (2017): 178-99.

Mogensen, Andreas L. “Staking Our Future: Deontic Longtermism and the Non-Identity Problem.” Manuscript, Global Health Priorities Institute, Oxford University.

Nebel, Jacob M. “The Good, the Bad, and the Transitivity of Better Than.” *Noûs* 52, no. 4 (2018): 874-99.

\*Qizilbash, Mozaffar. “On Parity and the Intuition of Neutrality.” *Economics & Philosophy* 34, no. 1 (2018): 87-108.

Schulte, Peter. “Von Glückspilzen und Unglücksrabben: Eine Untersuchung logisch-normativer Theodizeeprobleme.” In *Glaube und Rationalität: Gibt es gute Gründe für den (A)theismus?*, edited by Romy Jaster and Peter Schulte, 131-150, 2019. (In German)  
[https://doi.org/10.30965/9783957437891\\_008](https://doi.org/10.30965/9783957437891_008)

Spears, Dean. “The Asymmetry of Population Ethics: Experimental Social Choice and Dual-Process Moral Reasoning.” (2019).

\*Wasserman, David. “Death or Disability: The Carmentis Machine and Decision-Making for Critically Ill Children by Dominic Wilkinson.” *Kennedy Institute of Ethics Journal* 27, no. 1 (2017): E-4-E-11.

### Theses and Working Papers

Anderson, Matthew Lee. *In Defence of Children: Pro-and Anti-Natalist Arguments in Moral Philosophy and Karl Barth*. D.Phil. Thesis, University of Oxford, 2018.

Carlsmith, Joseph. “Helping People, Creating People, and Shaping the Far Future.” Manuscript, New York University, 2017.

Kaczmarek, Patrick Krystof. *A Fairness-Based Astronomical Waste Argument*. Ph.D. Thesis, University of Glasgow, 2018.

Kern, Adam B. *Tolerating the Terrible: Censorship, Speech, and Foreseeable Harm*. B.Phil. Thesis, Oxford University, UK, 2015.

Nebel, Jacob M. "Asymmetries in the Value of Existence." Manuscript, 2019.  
<https://jakenebel.com/papers/existencev4.pdf>

**"Future Persons and Victimless Wrongdoing" in Markus Rütter and Sebastian Muders (eds.) *Worauf es ankommt: Derek Parfitts praktische Philosophie in der Diskussion*. Hamburg: Felix Meiner Verlag, 2017.**

Published

Shelby, Tommie. *Dark ghettos: Injustice, dissent, and reform*. Harvard University Press, 2016.

Wallace, R.J., *The Moral Nexus* (Princeton: Princeton University Press, 2019).

**Conditional Reasons and the Procreation Asymmetry" (R&R at *Philosophy & Public Affairs*)**

Published

Pummer, Theron. "The Worseness of Nonexistence." In *Saving People from the Harm of Death*, edited by Espen Gamlund and Carl Tollef Solberg. Oxford University Press, 2019.  
DOI:10.1093/oso/9780190921415.003.0016

Shelby, Tommie. *Dark ghettos: Injustice, dissent, and reform*. Harvard University Press, 2016.

Wallace, R.J., *The Moral Nexus* (Princeton: Princeton University Press, 2019).

Working Paper

Budolfson, Mark, and Dean Spears. "Why the Repugnant Conclusion is Inescapable."

Manuscript, Climate Futures Initiative, Princeton University, 2018.

[https://scholar.princeton.edu/sites/default/files/cfi/files/budolfson\\_spears\\_2018\\_repugnant\\_cfi.pdf](https://scholar.princeton.edu/sites/default/files/cfi/files/budolfson_spears_2018_repugnant_cfi.pdf)

Campbell, Tim. "Personal Identity and Impersonal Ethics." Manuscript.

[https://www.academia.edu/40589926/Personal\\_Identity\\_and\\_Impersonal\\_Ethics](https://www.academia.edu/40589926/Personal_Identity_and_Impersonal_Ethics)

### III. CONTRACTUALISM AND INTERPERSONAL JUSTIFICATION

**“What We Owe to Hypocrites: Contractualism and the Speaker-Relativity of Justification”, *Philosophy & Public Affairs* 44.4 (2016): 223-265.**

Published or forthcoming

Crisp, Roger, and Theron Pummer. “Effective Justice.” Forthcoming in *The Journal of Moral Philosophy*, 2020. <https://philarchive.org/archive/CRIEJ-2v1>

Cross, Ben. “Radicalising Realist Legitimacy”, *Philosophy & Social Criticism* (2019).

Devolder, Katrien. “Euthanasia for Detainees in Belgium The Case of Frank Van Den Bleeken,” *Cambridge Quarterly of Healthcare Ethics* 25, no. 3 (July 2016): 384-394.

Dougherty, Tom. “Why Does Duress Undermine Consent?”, *Noûs* (2019).

de Kenessey, Brendan, “Promises as Proposals in Joint Practical Deliberation”, *Noûs* (2018).

Lippert-Rasmussen, Kasper. *Relational Egalitarianism: Living as Equals*. Cambridge University Press, 2018.

\*Mardellat, Victor. “Contractualism and the Paradox of Deontology.” *Philosophical Studies*, January 7, 2020. <https://doi.org/10.1007/s11098-019-01406-w>.

Walen, Alec. *The Mechanics of Claims and Permissible Killing in War*. Oxford University Press, 2019.

Theses and Working papers

Carlsmith, Joseph. *Hypocrisy and Accountability*. B.Phil. Thesis, Oxford University, 2017.

\*Cross, Ben. “Free Speech, Coercion, and Speaker-Relativity.” Manuscript. [https://www.academia.edu/34274446/Free\\_Speech\\_Coercion\\_and\\_Speaker-Relativity](https://www.academia.edu/34274446/Free_Speech_Coercion_and_Speaker-Relativity)

\*Scanlon, T.M., “Contractualism and Justification.” Manuscript, 2017. (To appear in the proceedings of the 2016 Lauener Conference.) [https://www.academia.edu/35861169/Contractualism\\_and\\_Justification.pdf](https://www.academia.edu/35861169/Contractualism_and_Justification.pdf)