

## How (Not) to Make Trade-Offs Between Health and Other Goods

Antti Kauppinen

University of Helsinki

Penultimate draft, July 7, 2021

Forthcoming in *Cambridge Quarterly of Healthcare Ethics*.

Pandemics are terrible not just because they cause suffering and death but also because measures to contain them have significant negative side effects on people's lives.

Interventions to prevent the spread of disease inevitably result in lost opportunities to work, learn, consume, and socialize, which tends to hurt most those who are already in a vulnerable position. There thus appears to be some kind of trade-off between protecting the health of at-risk people and the economic and other flourishing of others, at least in the short term.

Evidently, we shouldn't protect some people's health at *any* cost to others. How, then, should we assess such trade-offs from an ethical perspective?

It has been common to approach this issue as a question of *cost-effectiveness* of nonpharmaceutical health interventions (NPIs) in terms of some currency or another: which measure results in the best balance when we add up the gains and losses? (e.g. Singer and Plant 2020) My aim in this paper is to explore an alternative way to make principled trade-offs. Its roots lie in a conception of morality according to which equal respect towards individuals requires *impartial* consideration of their claims, not *impersonal* maximization of benefit (Rawls 1971, Nagel 1979, Scanlon 1998). An appealing way to flesh out the notions of impartiality and equal respect in this context is to adopt what I'll call the Justifiability Constraint: interpersonal trade-offs must be respectfully justifiable to their losers. Respectful justification requires that at least some on the winning side have clearly stronger claims to be

favoured than the losers, or that the contextually morally relevant claims on their side balance out the claims of the losers, taking the strength and number of both into account.

I'll make the case as follows. First, I introduce two cost effectiveness -based proposals, which serve as a foil for my positive view. The first is based on the most common measure of health outcomes, quality-adjusted life-years or QALYs. The thought is simple: NPIs will have a (roughly) predictable economic cost and impact on health outcomes, so we can calculate how many QALYs per pound (etc.) they are likely to produce. This method can be used not only to compare the cost-effectiveness of the measures in terms of health-related quality of life, but also to rule out those whose cost per additional QALY exceeds some fixed amount, as is standardly done in publicly funded health care systems like the UK NHS. After all, at some point the money saved can produce more good elsewhere. Many have proposed that we should adopt the same or similar standard for incremental cost-effectiveness of NPIs as we do for pharmaceutical interventions within the health care system (e.g. Miles, Stedman, and Heald 2020, van den Broek-Altenburg and Atherly 2020, and Vaillancourt 2020). I'll label this the generalized QALY or GQALY approach. The second proposal I look at generalizes the use of QALYs in a different way, and holds that government interventions should be chosen on the basis of cost-effectiveness in terms of *subjective well-being* or life satisfaction, as measured in wellbeing-adjusted life-years or WELLBYs (Frijters et al. 2020, Singer and Plant 2020, Frijters 2020). The simple, utilitarianism-inspired proposal is that NPIs like lockdowns should only be undertaken if they produce more WELLBYs than they cost.

Second, I argue that these proposals violate the Justifiability Constraint for the very reason they allow for cost-effectiveness calculations of NPIs: as aggregative outcome measures, they ignore facts about *which kind* of harm or benefit *different* individuals suffer or gain. Because of this, they count also (prospects of) benefits and harms that are in context

*morally irrelevant*, like the loss of a Christmas bonus or a minor reduction in health, which carry no weight against the loss of a life worth living. Such irrelevant harms could not be used to respectfully justify the choice to losers, even if aggregated benefits exceed aggregated harms (Kamm 1996, Voorhoeve 2017). Any single currency that enables trade-offs between any kind of benefit or harm whatsoever makes them *too easy* from an ethical perspective that takes the separateness of persons seriously.

Third, I outline how benefits and costs could be balanced against each other in the context of a pandemic without a common currency. On the impartial conception of morality, we need to balance all (contextually) relevant claims of individuals. How? I defend two basic principles, building on recent work in the contractualist tradition (Tadros 2019, Steuwer 2021a). The first, Always Count Relevant Claims, tells us to start by balancing the weakest claims for competing measures against each other, and then work our way towards the strongest ones by always weighing the remaining unbalanced claims against (possibly stronger) relevant ones. The second principle, No Tipping, says that though in this method weak claims can make a difference, they cannot alone tip the balance if they are irrelevant to the strongest opposing claim, because that could not be justified to the losers of the trade-off. I'll illustrate the use of these principles in some detail below.

Deliberation governed by these two principles requires not only relatively detailed information about what kinds of benefits and harms to different individuals are likely to result from alternative measures to contain the pandemic, but also fine-grained judgment about the relative strengths of moral claims grounded in prospects of such benefits and harms. Consequently, it falls far short of an algorithm that could be applied by a technocrat. Nevertheless, it's the best we can do when faced with heterogeneous claims of different individuals who merit equal respect in a high-stakes situation.

## 1 The Generalized QALY and WELLBY Approaches to Trade-Offs

As I noted at the start, effective measures to curb the spread of a contagious disease require isolating people from physical contact in various ways, which results in various costs as well as health benefits. I'm going to set aside the issue of whether, say, restrictions on free movement violate people's *rights*.<sup>1</sup> Whether there are such moral harms or not, it is very likely that lockdown measures have both readily quantifiable economic costs and hard-to-quantify adverse effects on many people's well-being. Estimates of these costs often fail to distinguish between the costs resulting from individuals' rational adaptation to the pandemic and the costs resulting from social distancing measures, thus significantly exaggerating the latter. In any case, stay-at-home orders, business closures, bans on public events and use of public spaces, and other similar measures will no doubt globally reduce economic activity on a scale of billions of dollars in any pandemic, which in turn reduces consumption and causes unemployment and poverty. They also contribute to an increase in domestic violence, loneliness, and even suicide.

Given that disease prevention measures thus have significant costs as well as benefits, there is a real need to find a way to weigh the two against each other in the case of each proposed measure. The natural thought is that "making trade-offs requires converting different outcomes into a single unit of value", as Peter Singer and Michael Plant (2020) put it. What, then, should be the 'single unit of value'? The two proposals I'll look at focus on health-related quality of life and well-being.

Start with health-related quality of life. For simplicity, I will here focus simply on quality-adjusted life years or QALYs, setting aside competing models like disability-adjusted life years. The idea of a QALY is simple: leading a full year in perfect health constitutes 1 QALY, while a health state no better than death gets you 0 QALYs. Based on population

---

<sup>1</sup> In my view, many such objections are based on conflation of legal rights with moral rights, which do not include the right to impose risk of serious harm on innocent others.

surveys, proponents of the approach have quantified values somewhere in between for many possible health states – for example, a year of life on dialysis is worth 0.75 QALYs.

QALYs can be and have been put to many uses. The most obvious is comparing the cost-effectiveness of treatments. Consider the following scenario:

*Cancer Choice*

Cancer treatment A costs £3,000,000 and is expected to yield additional 90 QALYs total for six patients, while treatment B costs £3,600,000 and is expected to yield 120 QALYs for the same number of different patients, or for a larger number of patients suffering from the same disease.

Let's assume that we're dealing with a public system and each patient has an equal claim to treatment. Which treatment if either should the hospital provide, if it can't offer both? With these figures, it's easy to calculate that treatment A produces an additional QALY for £33,333 and treatment B for around £30,000. Treatment B thus promises more value (understood as health-related quality of life) for money, so there is at least a prima facie case for choosing it (for complications that I'll ignore here, see Harris 1987 and Hausman 2010).

The second relevant use for QALYs is that they allow health care systems to set a *limit* to how much of their resources they will put into a particular use. Since health budgets are always limited, interventions that are costly but ineffective have significant opportunity costs in terms of health. No public system should spend a million to prevent ten men from losing their hair, since it would mean smaller resources for more serious medical problems, resulting in more preventable suffering. In the UK, the NHS normally considers a new treatment sufficiently incrementally cost-effective if it produces an additional QALY for less than £20,000 to 30,000, so at most treatment B above would be offered in the UK.

What I'm calling the Generalized QALY approach or GQALY says that we should assess the cost-effectiveness of nonpharmaceutical interventions in the same way as pharmaceutical ones. The simplest way to do this would be to hold that we should adopt the measure (which might be doing business as usual) that promotes health-related quality of life at the lowest cost per QALY. But this would leave open the question of how much health-related quality of life to buy at the low price in terms of other goods (maybe paying a little more per QALY would result in a much bigger total). Perhaps because of this, defenders of GQALY tend to focus on maximal acceptable incremental cost, as in the following:

*The Core Claim of GQALY*

If the opportunity cost of an additional QALY saved by a nonpharmaceutical intervention under consideration exceeds the maximal monetary cost per additional QALY considered acceptable for a health care intervention (even in an emergency situation), the measure should not be adopted.

It may be obvious that this use of QALYs goes against their original purposes, but the explanation of why this is the case turns out to be philosophically interesting. To begin with, consider the following case:

*Restaurant Closure*

Closing down restaurants in some part of the UK costs £10,000,000 a week and is estimated to save (each week) the lives of 30 old people, who would each live in perfect health for 5 years, thus resulting in 150 QALYs.

The cost per QALY of closing down is thus  $£10,000,000/150 = £66,000$ . This far exceeds the maximum that the NHS is ordinarily willing to pay for an additional QALY, so GQALY says

that under these assumptions, the restaurants should not be closed. Even if should pay more in an emergency situation, with suitable numbers we'll reach the limit at some point.

(For simplicity, this calculation ignores other benefits and costs of keeping the restaurants open, such as the pleasure people take in dining together, and the suffering of people who get ill but recover.)

QALY has been at least implicit in many different commentaries on COVID-19 policy. Health care economists come closest to explicit statements. For example, Miles et al. (2020) calculate the economic cost of UK restrictions up to June 2020 and compare them against the benefits as follows:

One approach is to focus on quality adjusted life years (QALYs) that may have been saved as a result of restrictions that have been in place in the UK up to the end of June and to convert that to a metric that can be compared with estimates of the cost of the restrictions. [...] We will assume that the benefit of the restrictions that prevented such deaths are the value of ten quality-adjusted years of life multiplied by the number of lives saved. The NICE £30,000 threshold is an assessment of the (maximum) resource cost that would be justified for the UK health service to make an expected saving of one quality adjusted year of life. To save ten QALY would be worth up to £300,000. [Given the economic costs,] the cost per QALY saved of the lockdown looks to be far in excess (generally by a factor of at least 3 and often by a factor of 10 and more) of that considered acceptable for health treatments in the UK. (9, 11, 13)

Many more examples can be found in online commentary on COVID-19 measures by economists (e.g. van den Broek-Altenburg and Atherly 2020, Teather 2020, Young 2020, Vaillancourt 2020). Philosophers like Julian Savulescu et al. cite such reasoning with

apparent approval, noting of the initial US stimulus package that “such measures are unlikely to be cost-effective according to the usual thresholds applied to the costs of medical interventions to save lives.” (2020, 624)

### *From QALYs to WELLBYs*

A different way around the problem of figuring out how much to pay for health-related quality of life at the expense of other goods is to look at what the total impact of proposed measures is on people’s overall well-being. Perhaps we should simply impose pandemic restrictions to the extent that doing so maximizes net well-being (e.g. Savulescu et al. 2020). Some might think, to be sure, that measuring overall well-being is a very different task than just looking at health-related quality of life. But proponents of recently popular science of well-being think it is not so difficult. In particular, there exists a burgeoning literature on subjective well-being (SWB), measured in various ways (see Diener et al. 1999). The simplest and most widely used currency is life satisfaction, which is measured in the UK by asking the question “‘Overall, how satisfied are you with your life nowadays’, where a 0 is ‘not at all’ and a 10 is ‘completely’” (Frijters et al. 2020, 16) Among others, Paul Frijters and co-authors (2020) make use of life satisfaction measures to define a wellbeing-adjusted life-year, or WELLBY. A year lived at life satisfaction level 1 on a scale of 1 to 10 yields 1 WELLBY; thus, given that the average life satisfaction level in rich countries tends to be around 7.5, their average citizens gain 7.5 WELLBYs each year.

For proponents of the WELLBY approach, then, subjective well-being is the currency in terms of which trade-offs between different goods should be made. (Let us accept (purely!) for the sake of argument the identification of life satisfaction with well-being and any measurement assumptions WELLBY proponents make.) As Frijters et al. 2020 say, “Wellbeing-adjusted life-years (WELLBYs) expand the QALY idea to the whole of life, not



just health. It is hence a tool for judging the trade-offs between large government programmes” (4). In Layard et al. 2020, WELLBYs are applied to assessing lockdown measures in the UK, in particular addressing the question of what is the best date for lifting them in order to maximize WELLBYs. To convert various observed effects of lockdown and the virus, the authors make use of findings from the science of happiness. For example, when it comes to money, they report a robust finding that “a 1% gain in income increases wellbeing (measured 0-10) by around 0.002 points” (Layard et al. 2020, 5) The psychological cost of being unemployed for a year, in turn, is 0.7 WELLBYs.

Making use of such figures, Frijters (2020a) calculates that even in the best case scenario, lockdown-produced unemployment alone will globally cause the loss of 50 million WELLBYs in 2020-2021, which, according to him, is equivalent to the harm caused by 1.7 million deaths of people with 5 good years left. Taking other well-being costs into account as well, he concludes that no restrictions should ever have been introduced, even if this resulted globally in 200 million deaths of people with three years left (Frijters 2020b).

## **2 Aggregation, Justifiability, and Balancing Competing Claims**

It certainly makes sense to look for a common currency in terms of which to make trade-offs between different harms and benefits when we can’t have everything we want. But this effort comes with dangers that may not be immediately apparent, and I’m going to argue that both GQALY and WELLBY approaches end up endorsing unjustifiable trade-offs because of it.

The source of the trouble is that what is at issue are *interpersonal* trade-offs, in which different choices result in *different harms* to *different people*. To be sure, there are approaches to ethics, such as Benthamite quantitative utilitarianism, according to which all that matters for right action is how much well-being alternative options bring about, so that it makes no difference *who* is harmed or benefited and how much. On this view people are, as

Tom Regan (1983, 205–211) says, merely ‘receptacles’ of intrinsic value who are in themselves bereft of non-derivative moral significance – all that matters is achieving the best *impersonal* sum of benefits and harms, not where we ‘pour’ them. I’m not going to argue directly against this radically (and openly) counterintuitive picture of ethics (for classic arguments, see e.g. Rawls 1971 and Nozick 1974), but will instead spell out the consequences of a fundamentally different conception in the hope that their independent plausibility will suffice to raise further trouble for the impersonal conception.

On the alternative picture, it is *persons* who have fundamental moral significance, and their well- or ill-being matters for morality because it matters to them. In the notoriously ambiguous Kantian phrase, they are ends in themselves, not mere means for maximizing utility. Like the impersonal picture, this alternative holds that everyone counts equally, but interprets this in terms of impartial consideration of the morally relevant claims of individuals rather than adding up benefits and harms regardless of where they occur, so I will call it the *impartial conception of morality*. There are many different ways of specifying what impartiality between persons amounts to, but the core idea is that everyone’s personal and partial perspective carries equal weight from the moral perspective. In some way, morality incorporates the partial viewpoints of individuals rather than overriding them (Nagel 1979). The thin idea that I’ll rely on here is that equal respect requires the principles behind choices to be justifiable to all, insofar as they are reasonable, while allowing everyone some partiality to themselves. This is weaker than a contractualist view, which would hold, very roughly, that acts are right *because* principles allowing them couldn’t be reasonably rejected (Scanlon 1998), so it is compatible with many different forms of nonconsequentialism and sophisticated consequentialism (Portmore 2011).

In the particular case of interpersonal tradeoffs, I’m going to assume the following simple principle:

*The Justifiability Constraint*

Principles for interpersonal trade-offs must be justifiable to their losers, insofar as they are reasonable.

As a shorthand, I'll also say that the principles must be *respectfully* justifiable (see Steuwer 2021a). In keeping with my ecumenical aims, I will not attempt to specify exactly when it is reasonable for a loser to reject a principle for interpersonal trade-offs, but the basic idea is that they must have sufficient reason to do so without giving excessive preference to their own self-interest. Thus, a principle can be justifiable to a loser even if she does not, in fact, accept it.

Because on the impartial conception of morality it is equal consideration of the morally relevant claims of each individual that matters, it sometimes rules out impersonal maximization of well-being. In Thomas Scanlon's Transmitter Room scenario, the choice is between a worker caught in the TV transmission system suffering fifteen minutes of extreme agony and a billion people losing live view of the last fifteen minutes of the World Cup final (1998, 235). Scanlon's view is that though the small though non-trivial benefits to the viewers taken together add up to a total sum of well-being that far exceeds disvalue of the worker's suffering, continuing the broadcast is not respectfully justifiable to the worker, since his claim to have it stopped is far weightier than any other individual's claim to keep it going.

At the same time, however, accepting the Justifiability Constraint doesn't by itself rule out taking the number of affected individuals into account when impartially weighing claims. For example, it may be unreasonable for me to object to a principle that imposes a burden on me just because a larger number of others would otherwise suffer from a similar burden (Scanlon ms). That's why the Justifiability Constraint doesn't simply beg the question

against the use of QALYs, for example – at least sometimes, ‘better value for money’ does justify a trade-off.

Consider Cancer Choice. How could the choice of treatment B be justified to the patients who thereby lose out on treatment A? Clearly, the answer must be that the limited resources are put to better use – same number of patients in equal need get more life-years as a result, or more patients get the same benefit. We can say to patients in the A group: “We’re sorry that we can’t offer you treatment, but the alternative would have been a greater loss to an equal number of people or an equal loss to a greater number of people – we couldn’t have justified *that* choice to them.” Importantly, a similar justification is sometimes available for people who don’t receive treatment, because the incremental cost of providing it would exceed a cost-effectiveness limit like the NHS one: offering the treatment would require foregoing a greater benefit to others in similar need, given budget constraints.<sup>2</sup>

What about Restaurant Closure, then? Here, instead of taking action as a result of which 30 people a week would avoid getting sick and dying, the QALY approach tells us that we should keep the restaurants open to avoid losses of £10,000,000 a week. How can we justify this choice to the 30 people who are about to lose their lives as a result? For the justification to be parallel to the cancer treatment case, we’d have to say that the money saved will be put to better use elsewhere, to meet equally urgent needs. But that’s precisely what we *can’t* say here, when all we know is that the keeping the restaurants open will generate £10,000,000. Some of that money will be income that goes to the restaurant owners, managers, and workers, as well as the public purse via tax, and circulates back via consumption to feed further economic activity. Much of it could go in a safe, or be burned in an art performance, or buy another Bentley for a restaurant owner; some would pay for a

---

<sup>2</sup> This is not to say that using QALYs in health care choices in general meets the Justification Constraint. I will set aside this contentious issue here.

dentist, some a birthday present to a child. All we can say for sure (more or less) is that it will be used to *satisfy some people's preferences*, whatever they may be.

And this, I claim, is not a sufficient justification for the significant loss imposed by the continued economic activity in pandemic conditions. Specifically, this is because some preference-satisfactions are *morally irrelevant utilities* in the context of life-and-death choices. By saying that some utilities are morally irrelevant I mean that they carry no weight in the context of the decision at hand, given the kind of reasons there are on the other side, which I'll argue makes a difference to how we should make the trade-off.

Perhaps the best way to illustrate the notion of relevance is to start with Frances Kamm's Sore Throat Case (Kamm 1993, 165). As she originally puts it, we have some scarce medicine that could save either Jim or Joe, neither of whom has any special claim to it and both of whom have an equal claim to our assistance. The right thing to do is to flip a coin. But now suppose that if and only if the medicine is given to Jim, there is enough left over to cure a sore throat that Nancy will otherwise have next week. Now, giving the medicine to Jim would result in an outcome that is strictly better than the outcome of giving it to Joe. Nevertheless, Kamm says, it would be morally wrong to give the medicine to Jim on this basis. Her grounds for saying this are, roughly, that from Joe's perspective, it would be as if something (Nancy's sore throat) that would not be a contender in a pairwise comparison with saving his life defeats saving his life, and (in keeping with the impartial conception) that an objective moral perspective should incorporate subjects' own perspectives. We would be "showing insufficient concern for what Joe's chance to live means to him" if we allowed the utility of curing the sore throat count against it (1993, 195-196). It would not be respectful. In other choice contexts, to be sure, Nancy's sore throat would be a morally relevant consideration (suppose the choice is between curing Nancy's sore throat and having a beer). For moral relevance, it matters what's on the other side of the scales.

What makes a benefit or harm contextually morally relevant and why? Evidently, other things being equal, opposing claims that are based on *identical* benefit or harms are relevant to each other from an impartial perspective. But what about when the outcomes for different people are not identical? Plausibly, the strength of one's claims on others depends *inter alia* on the amount of harm one would suffer if not favoured and the size of benefit one would gain if favoured (e.g. Kamm 2007, 57). The strength of one's claims also depends on *how* the harm or benefit is brought about – famously, we have a stronger claim against being harmed as a means to benefit others than against being harmed as a side effect of benefiting others, and a stronger claim against someone doing harm to us than merely allowing us to suffer harm they could have prevented. For reasons of space, I'll have to largely set aside the latter kind of issues here.

When do some people's weaker claims count against the stronger claims of others? The most promising proposals draw on the idea that A's weaker claim is relevant to B's stronger opposing claim if and only if *morally permissible self-concern* allows A to prefer an outcome in which her claim is satisfied over the satisfaction of B's claim (Voorhoeve 2014, 71-72). This is a plausible way of cashing out the idea that an impartial perspective incorporates individual perspectives rather than merging them – if it's morally permissible for A to prefer that she keeps an arm to B keeping both arms, then surely A's loss of an arm is *relevant* from an impartial perspective as well, even if it is outweighed by B's greater loss in a one-to-one comparison. As Alex Voorhoeve (2017) puts it, an impartial spectator can sympathize with A pressing her claim against B's. In contrast, an impartial spectator could not sympathize with Nancy pressing her claim to get the medicine for her sore throat when Joe's opposing claim is based on his survival. Nor could Nancy complain about a third party giving the medicine to Joe when it wouldn't be morally permissible for herself to give so much preference to herself, if *she* were the one making the choice (Voorhoeve 2014, 74).

The reason why relevance matters in this context is that the most plausible principles for restricted aggregation say that impartial concern requires *balancing all relevant claims* (BARC) (cf. Scanlon 1998, 239–40, Voorhoeve 2014, Tadros 2019, Steuwer 2021a). The basic rationale is that while it would not be compatible with equal respect to fail to save someone because of many other people’s irrelevant claims (as in Transmitter Room), equal respect *does* require giving weight to each person’s *relevant* claims. For example, in a choice between curing my arthritis or two other people’s arthritis, the benefit to me is balanced by the benefit to one of them, and the remaining person’s relevant and unmatched claim tips the balance in their favour (Kamm 2007).

To illustrate the implications of balancing all relevant claims for evaluating NPI tradeoffs, let’s first consider a couple of *highly* simplified and unrealistic scenarios (assume that all other morally relevant facts are the same regardless of one’s choice). The first involves a significant harm, losing a home:

Close businesses	Keep businesses open
No elderly die	10 elderly die (as a side effect)
10000 people become homeless	Nobody loses their home

Plausibly, becoming homeless (and thus having to live in an extremely vulnerable position, subject to many hazards) is a source of a relevant claim even against the claim of someone who would die several years prematurely as a foreseen but unintended side effect. BARC then entails that when a sufficient number of people are made homeless by closing businesses, they should not be closed, even if it means that some elderly people will predictably lose their lives in consequence. It would not be reasonable for the elderly to reject this trade-off, given the seriousness of homelessness and the number of those affected, so it is

justifiable to them (even though they could obviously reasonably reject being killed as a means to save others). Contrast this with the following:

Close businesses	Keep businesses open
No elderly die	10 elderly die
100000 people lose Christmas bonuses	Nobody loses their Christmas bonuses

In this case, it's plausibly not within the bounds of permissible self-concern to prefer keeping one's Christmas bonus to another losing a life. Thus, if we impartially balance the moral claims that each individual can legitimately make on others, we won't take the Christmas bonuses into account, and the strong claim of the elderly is unopposed.

It seems, then, that balancing relevant claims gives the right results for these simple cases. But in more complex scenarios, various challenges arise. I will set aside here concerns about intransitivity of choiceworthiness and violating the Principle of Irrelevant Alternatives<sup>3</sup>, and focus on a recently influential line of objection based on mixed groups, since mixed groups are the norm in real-world situations. For reasons of space, I'll only consider a two-stage case inspired by Patrick Tomlin (2017; cf. Horton 2020). Let's start with the following scenario, in which we'll stipulate it's permissible to flip a coin to decide whether to close businesses, since the claims on either side are balanced:

Close businesses	Keep businesses open
100 homeless	10 elderly die

At stage 2, further effects are added:

Close businesses	Keep businesses open
100 homeless	10 elderly die
1000 people lose job for 3 months	1 person loses their job for 3 months

---

<sup>3</sup> For these worries and good replies, see Parfit 2003, Kamm 2007, and Voorhoeve 2014.



Intuitively, adding many more suffering the same harm on the side of closing businesses means that it is still at least permissible keep businesses open. But suppose that losing a job for 3 months is relevant in comparison to homelessness, but not to death. Early defenders of restricted aggregation, like Voorhoeve (2014), thought of relevance as an on/off matter decided by the strongest competing claim. On such a view, the 1000 job losses when businesses are closed wouldn't count, because they compete with death. The single job lost when businesses are open *would* count, however, because it competes only with homelessness, and would thus tip the balance in favour of closing businesses. But, Tomlin objects, surely this goes against the plausible idea that equal claims should get equal consideration – after all, we added far more claims of the same kind *against* closing businesses! (2017, 241)

Fortunately, this is not a problem for BARC, if we adopt two simple and well-motivated principles. The first, Always Count Relevant Claims, says that all unbalanced claims always count against all relevant claims (cf. Tadros 2019) – after all, how could an impartial approach deny that? To ensure this, we should start balancing from the weakest claims and move towards the strongest. Second, no claims irrelevant to the strongest opposing claim can tip the balance (No Irrelevant Tipping), because otherwise the choice could not be respectfully justified to the biggest losers (Steuwer 2021a, 26). In the two-stage case, Always Count Relevant Claims means that the thousand job losses when businesses are closed easily beat the one job loss if they're open, so the balance doesn't tip towards closing businesses by the addition. At the same time, No Irrelevant Tipping ensures that adding the many short-term jobless on the side of business closure does not break the tie in favour of keeping businesses open – it would not allow for a respectful justification to those who would consequently die. (In the next section, I'll illustrate in more detail how these principles work, and discuss the role of uncertainty.)

While there is a lot more to say about this ongoing theoretical debate<sup>4</sup>, I believe that at the end of the day, the thought that some but not all harms count against others from a moral perspective is very appealing. And that spells trouble for GQALY, because, as they say, a dollar is a dollar. Once we start making trade-offs between goods, say health and the economy, *outside the healthcare context* on the basis of any monetary value whatsoever for a QALY, we *are* counting any benefit that money can buy or harm that money can prevent against lost lives.<sup>5</sup> While some of these harms – say, homelessness or loss of access to higher education – will indeed be relevant in comparison to deaths, we’ve lost the means to distinguish them from others. There will be situations in which irrelevant claims tip the scale, and the trade-off can’t be respectfully justified to its losers.

What about the WELLBY approach, then? Consider the claim by Paul Frijters (2020a) that 80 million people unemployed for five years is equivalent to at least 1.7 million deaths of people with five good years left in terms of WELLBYs lost. Let’s accept these figures for the sake of argument. Dividing through and approximating, we get the result that if we have to choose between one currently healthy person dying now rather than in five years and four people losing their job for the same number of years, we should choose the option that leads to the death of the one person now. While unemployment is no doubt a serious social ill – though just how serious depends on whether the unemployed have access to decent housing, health care, and so on – it is difficult to see how such a choice could be defensible to the person about to lose her life. After all, she stands to lose much more individually than any of those who would otherwise lose their jobs. Provided that the

---

<sup>4</sup> For example, what I say here does not yet address Horton’s (2018) dilemma for limited aggregation, based on the seemingly counterintuitive result that for BARC, verdicts about subgroups do not agglomerate. I hope to return to this elsewhere (though Tadros 2019 and Steuwer 2021a already give persuasive responses).

<sup>5</sup> What is more, we are counting them in the same way, regardless of whether they mean an extra boost to someone who is well-off or stand between utter misery and tolerable existence. This gives rise to a *prioritarian* objection to GQALY based on the idea that losses to the worse-off should count for more – in the terms I’m using, the worse-off can reasonably reject a principle granting the same value to equal benefits for the better-off. For reasons of space, I can’t pursue this here.

unemployed are not mired in misery, an impartial spectator could not sympathize with their objection (cf. Kauppinen 2018).

So I conclude that the WELLBY approach does no better than GQALY, and for the same reason – using a single currency to enable trade-offs between any goods and bads whatsoever makes trade-offs too easy, forcing us to make trades that we shouldn't make.

### **3 Balancing Claims in a Pandemic**

Suppose I'm right, and using any common currency will result in ignoring morally relevant differences between harms and causal pathways. How, then, should we assess the inevitable trade-offs between health and other goods? Here is a rough sketch of what the procedure would need to look like on the basis of the above considerations:

1. For each proposed measure, identify and model the sources of claims that individuals on either side of the trade-off have for favouring them.
2. Estimate the strength of each claim relative to others and sort them into relevance categories (e.g. weak, moderate, and strong claims). The strength of a claim depends not only on the amount of benefit or harm to the individual, but also on how the benefit or harm would be brought about (and possibly on their *ex ante* likelihood – see below).
3. Balance the mutually relevant claims on both sides in accordance with Always Count Relevant Claims and No Tipping, taking into account both strength and number. Here's how to proceed: first balance weak claims against each other; weigh any remaining unbalanced weak aims against opposing moderate claims; then weigh remaining unbalanced moderate claims against each other; weigh any remaining unbalanced moderate claims against opposing strong claims; finally, weigh the remaining unbalanced strong claims against each other. At the end, make sure that the

balance doesn't ultimately tip against the strongest claims only because of weak opposing claims.

4. Choose the policy favoured by the balance of relevant claims after pairwise comparison of all options, because it will be respectfully justifiable to the losers of the trade-off.

To illustrate how this method works, consider first having to decide between the following policies:

A (Do little)	B (Close businesses)
15 lose last 5 years (strong claim against A)	10 lose last 5 years (strong claim against B)
1000 suffer mild illness (weak claim against A)	1000 lose their homes (moderate claim against B)

Following Always Count Relevant Claims, the weak claims against A are first weighed against moderate claims against B, balancing some of them, but leaving most of them to count against the claims of those who would die in A. On the assumption that homelessness is a serious enough harm to count against losing five years of life, the many moderate claims against B then balance some or all of the strong claims against A. Even if, say, five people still have a strong unbalanced claim against choosing A, once they are balanced with the strong claims against B, the result is that some strong claims against B are left unopposed. Option A is then the right one to take, even though it involves somewhat more of the most serious harms. Note that if instead of homelessness, B involved the loss of a million Christmas bonuses, which would ground only weak claims against it, the same method would result in B being the right policy, since as contextually irrelevant, the bonuses would not balance any of the excess deaths in A. (No Tipping doesn't come into play here.)

Before looking at more realistic cases, I must address an important simplifying assumption I've been making by talking as if the strength of people's claims against policies

depends on what actually ends up happening if they are adopted. In recent discussions of contractualism, this is called the *ex post* perspective. An evident problem for it is that it can easily end up prohibiting seemingly acceptable policies that are beneficial to many but involve a *risk* of serious harm for someone. For example, as Elizabeth Ashford (2003) observes, personal air travel inevitably imposes a risk not only on travelers themselves but on people who don't benefit from it in any way – say, a plane might fall on an Amish community. The burden borne by those who die in accidents grounds a stronger claim against allowing air travel than the burden that would-be travelers have to bear if it's prohibited, so *ex post* contractualism seems to prohibit the practice. Such scenarios have convinced many that the strength of people's claims against principles or policies depends instead on the *ex ante prospects* of benefit and harm (Suikkanen 2020). As Johann Frick puts it, "A person's harm-based complaint against a loss she suffers must [...] be discounted by her *ex ante* unlikelihood of suffering a loss and by her *ex ante* likelihood of benefiting from the risky action" (2015, 188). Clearly, if the Amish have, say, one in a million chance of being killed by a falling plane and their claim against allowing personal air travel is grounded in the tiny prospect of being killed rather than the badness of actually being killed, their complaint is easily outweighed by the prospective benefits of air travel to travelers, and the practice can be respectfully justified to them (including those who in the bad case end up being killed!).<sup>6</sup> To be sure, it's by no means obvious how to calculate the relevant probabilities (for example, the risks people face often depend on their own choices, which may be morally relevant), and simple discounting based on likelihood can lead to counterintuitive consequences (Otsuka 2015), so the debate between *ex ante* and *ex post* justifiability is far from settled.

---

<sup>6</sup> On Frick's view, if it is known in advance that *I*, for example, will die as a result of a vaccine that will save everyone else from a smaller harm that is irrelevant in comparison with death, mandating vaccination will *not* be justifiable, because my objection to it will not be *ex ante* discounted by the improbability of my death. Steuwer (2021b), in turn, rightly observes that the issue isn't really whether it is *known* who will die, but rather whether there is some specific person with, say, a genetic predisposition, who *will* die if the mandate is imposed. If that's the case, the doomed person will have an undiscounted claim against the mandate.

What would an *ex ante* perspective mean for the present proposal? In short, it would complicate things, but the fundamentals wouldn't change. In the best case scenario, we might have reason to be 95% confident that unless some measure is adopted, people in an At-Risk group have an 8-10% chance of dying, while if it is adopted, people in a Safe group have a 25% chance of losing their job for six months. Suppose A is in the At-Risk group and B is in the Safe group. Clearly, *ex ante* A's claim to adopt the measure is weaker relative to B's claim to avoid it than it would be if it was certain that A would die without the measure, since it's discounted by the relatively small likelihood of dying. But it can still be much stronger than B's claim. And note that the more tightly we circumscribe the groups – maybe a specific At-Risk group is overweight elderly bus conductors with a heart problem – the worse the probability-weighted prospects are for some. Thus, there will still be some individuals for whom the chance of serious harm without NPIs is quite high, and whose *ex ante* claim for adopting the measures is correspondingly strong. It is akin to my claim against being subjected to Russian Roulette against my will, even if I don't end up being harmed.

For the purposes of this paper, I won't take a stand on whether we should assess the strength of reasons from an *ex ante* or *ex post* perspective when considering justifiability. What I say about balancing claims works for either option. The difference is that from the *ex ante* perspective, the antecedent chances of different possible outcomes influence the strength of people's claims. Consequently, the result of the comparison is more likely to lean towards avoiding smaller but likelier harms to many, since more serious but less probable harms are discounted more heavily.

To show, then, that complexities involved in taking account qualitative differences among harms and benefits do not make principled trade-offs impossible, let us look at a somewhat more realistic toy scenario for a population of 100 000, focusing on the outcomes at the peak of the probability distribution (which are the sources of the strongest *ex ante*

claims and likeliest *ex post* claims). I'll sort out sources of claims to different stringency categories, asking first which weaker claims it wouldn't be morally permissible to press against stronger ones in a one-to-one conflict to establish irrelevance, and then placing in the middle claims that are relevant to both strong and weak ones. Finally, I'll use a no-pandemic scenario as a baseline for comparison, bearing in mind that lack of restrictions also has non-health costs:

	No restrictions	Moderate restrictions (no large gatherings, limited restaurant hours, universities closed etc.)	Strict restrictions (stay-at-home orders, schools and shops closed etc.)
Sources of strong claims	+3000 <b>deaths in At-Risk population</b> (loss of at least 5 years of life worth living) +500 <b>destitute</b>	+600 deaths in At-Risk population +1500 destitute	+300 deaths in At-Risk population +3000 destitute
Sources of moderate claims	+1000 people with <b>permanent lung damage</b>	+2000 <b>significant mental health problems</b>	+6000 significant mental health problems +500 <b>children falling permanently behind in education</b>
Sources of weak claims	+4000 temporarily <b>unemployed</b> +4000 suffering temporary <b>loss of leisure opportunities</b> , +20 000 temporary <b>significant loss of life satisfaction</b> +25 000 <b>mild illnesses</b>	+8000 temporarily unemployed +15 000 temporary losses of leisure opportunities +25 000 temporary significant loss of life satisfaction +8000 mild illnesses	+13000 temporarily unemployed +60 000 temporary losses of leisure opportunities +35 000 temporary significant loss of life satisfaction +3000 mild illnesses

To simplify, as we go from loose to strict restrictions, there's fewer deaths and physical illnesses, but more poverty, mental health issues (like depression and severe loneliness), educational problems, unemployment, and losses of life satisfaction (reflecting lost consumption, leisure opportunities, socializing, and so on).

The next step, then, is pairwise comparison of options. This requires settling, for example, what number X of people rendered destitute gives rise to an equally urgent claim against restrictions as a certain number Y of lives otherwise lost gives for restrictions (or, alternatively, what risk of destitution grounds equal claims to some risk of death). I've argued that there is no common currency like QALYs in terms of which to make such comparisons, but that's not to say that the kind of valuation methods used to evaluate health-related quality of life, such as the standard gamble and the time trade-off method (Bognar and Hirose 2014, 36-40), couldn't be useful here. For example, people would presumably give up some of their lifetime in order to avoid destitution (losing all property, living on the street), and that number is presumably lower than what they'd give up to avoid unemployment. Such information at least bears on justifiable trade-offs between the two kinds of harm.

If the numbers are as in the table above and what is listed there is representative of other morally relevant considerations, it seems that Moderate restrictions are favoured over No restrictions. Let's start with weak claims on both sides. I've stipulated a steep reduction in mild illnesses as a result of the Moderate restrictions, which should roughly balance the additional losses of leisure opportunity and life satisfaction. The claims of the extra 4000 temporarily unemployed because of the restrictions do still count against them, and need to be weighed against moderate claims next. They will balance some of the claims in favour of restrictions arising from permanent lung damage. Taking this into account, let's say there remain 1200 moderate claims against Moderate restrictions arising from mental health.

These moderate claims then join with the claims of the extra 1000 destitute people against restrictions to compete against the strong claims of the additional 2400 people in At-Risk groups who would lose at least five years of life worth living without restrictions. Or, to put it in *ex ante* language, against the strong claims of all members of the At-Risk group



based on the significant risk of losing their life.<sup>7</sup> (For simplicity, I'll set aside *ex ante* phrasing going forward.) Since these claims are mutually relevant, those of the depressed and destitute will balance some of the claims of those would lose their lives. Given that we are talking about severe and lasting hardship for several thousand people, my verdict is that such claims would suffice to respectfully justify measures resulting in several hundred people getting killed by the virus as a side effect. However, in this case, as the numbers stand, it seems to me that all these hardships do not strike out all the claims of 2400 extra people who would die.<sup>8</sup> So given these stipulations, it is possible to respectfully justify Moderate restrictions to those who lose as a result, while it wouldn't be possible to respectfully justify No restrictions to the losers of that option.

How about the comparison between Moderate and Strict restrictions? Starting again from weak claims, while people being largely confined to their homes reduces mild cases of illness by a further 5000, this is plausibly already outweighed by significantly reduced life satisfaction and leisure opportunities, leaving at least the claims of the extra 5000 unemployed to compete against moderate claims for Strict restrictions. As before, these claims balance some of the mental health issues. Let's say this leaves the claims of 4500 mental health sufferers and 500 children falling permanently behind in education to be balanced against the strong claims arising from the additional 300 lives saved by Strict restrictions. They will surely balance many of them, and very plausibly the claims of the additional 1500 made destitute by Strict restrictions more than balance the rest. If so, Moderate restrictions can be respectfully justified to those who would have done better under Strict restrictions, but not vice versa. Thus, overall, we have established by way of piecemeal

---

<sup>7</sup> If there are 30 000 people in the At-Risk group, it's easy to calculate that they face 1 in 10 odds with No Restrictions and 1 in 50 odds with Moderate Restrictions.

<sup>8</sup> To comply with No Tipping, we should ask at this point whether it's ultimately only because of weak opposing claims that the balance tips against No restrictions. The answer is negative, as evident from the fact that the balance of weak claims counts *for* them.

but principled balancing of relevant claims that Moderate restrictions are the most choiceworthy of the modeled options.<sup>9</sup>

To sum up, then, in this final section I've given a sketch of how we can reason about trade-offs between health and other goods without using a single currency in terms of which to make them. Balancing heterogeneous reasons no doubt results in a more complicated procedure, and carrying it out requires both judgment and much more granular information than QALY or well-being calculations. But that is only to be expected if we take seriously the moral conception of individuals and their diverse claims on each other mattering for their own sake. In the real world, resolving such trade-offs is much harder yet, since there are even more sources of relevant claims, and reliable modeling of all relevant expected consequences of appropriately fine-grained options is challenging and time-consuming. There is sure to be uncertainty about the facts and disagreement about the strength of the moral claims they ground.

However, we shouldn't exaggerate the extent of the disagreement either. People may grumble, but it is likely that few would be willing to insist that they should be allowed to go to the movies if it meant that other people would be likely to die as a result. Surveys using familiar methods for assessing first-person trade-offs and deliberative polling (Fishkin 2018) about interpersonal trade-offs would lend democratic legitimacy to judgment calls about weight and relevance. Laying out such principles in policy documents – for example, saying that the government assumes that the claims of a thousand people unemployed because of a policy more than balance the claims of a dozen depressed without it – would allow for a kind of *public* justification of interpersonal trade-offs to their losers that would in the best case coincide with their moral justification. What I've argued for here is that only the availability

---

<sup>9</sup> For completeness, we could also compare No and Strict restrictions, but that is not necessary here.

of such moral justification would ensure that even those who must bear the cost of mitigating a tragic situation are treated with the respect they deserve.<sup>1</sup>

## Sources

Ashford E. The demandingness of Scanlon's contractualism. *Ethics* 2003;113(2):273–302.

Bognar G, Hirose I. *The Ethics of Health Care Rationing*. London: Routledge; 2014.

van den Broek-Altenburg, E, Atherly A. Economic cost of flattening the curve. The Incidental Economist blog 2020 Mar 24; available at <https://theincidentaleconomist.com/wordpress/economic-cost-of-flattening-the-curve/>.

Conover C. How economists calculate the costs and benefits of COVID-19 lockdowns. *Forbes*, 2020 Mar 27; available at <https://www.forbes.com/sites/theapothecary/2020/03/27/how-economists-calculate-the-costs-and-benefits-of-covid-19-lockdowns/#633cee916f63>.

Diener E, Suh EM, Lucas RE, Smith HL. Subjective well-being: Three decades of progress. *Psychological bulletin* 1999;125(2):276–302.

Fishkin J. *Democracy When People Are Thinking*. New York: Oxford University Press; 2018.

Frick J. Contractualism and social risk. *Philosophy and Public Affairs* 2015;43(3): 175–223.

Frijters P. How many WELLBYs is the corona panic costing? Club Troppo blog, 2020a Apr 8; available at <https://clubtroppo.com.au/2020/04/08/how-many-wellbys-is-the-corona-panic-costing/>.

Frijters P. The corona dilemma. Club Troppo blog, 2020b Mar 21; available at <https://clubtroppo.com.au/2020/03/21/the-corona-dilemma/>.

Frijters P, Clark A, Krekel C, Layard R 2020. A happy choice: well-being as the goal of government. *Behavioral Public Policy*. doi:10.1017/bpp.2019.39.

Harris J. QALYfying the Value of a Life. *Journal of Medical Ethics* 1987;13:117–23.

- Hausman DM. *Valuing Health*. Oxford: Oxford University Press; 2018.
- Horton J. Always aggregate. *Philosophy and Public Affairs* 2018;46(2):160–74.
- Horton J. Partial aggregation in ethics. *Philosophy Compass* 2020;16:1–12.
- Kamm F. *Morality, Mortality Vol. 1. Death and Whom to Save From It*. New York: Oxford University Press; 1993.
- Kamm F. *Morality, Mortality, Volume II: Rights, Duties and Status*. New York: Oxford University Press; 1996.
- Kamm F. *Intricate Ethics*. New York: Oxford University Press; 2007.
- Kauppinen A. Sentimentalism, blameworthiness, and wrongdoing. In: Debes R, Stueber K, eds. *Ethical Sentimentalism*. Cambridge: Cambridge University Press; 2017:133–52.
- Layard R, Clark A, De Neve J-E, Krekel C, Fancourt D, Hey N, O'Donnell G. When to Release the Lockdown: A wellbeing framework for analysing costs and benefits. IZA DP No. 13186; 2020.
- McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics* 2008;26(9):733-44.
- Miles D, Stedman M, Heald A 2020. Living with Covid-19: Balancing costs against benefits in the face of the virus. *National Institute Economic Review* 2020 (253).
- Nagel T. *Mortal Questions*. Cambridge University Press; 1979.
- Parfit D. Justifiability to Each Person. *Ratio* 2003;16(4):369–90.
- Portmore D. *Commonsense Consequentialism*. New York: Oxford University Press; 2011.
- Regan T. *The Case for Animal Rights*. University of California Press; 1983.
- Scanlon T. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press; 1998.
- Scanlon T. Contractualism and Justification. ms.

- Singer P, Plant M. When will the pandemic cure be worse than the disease? *Project Syndicate* 2020 Apr 4. Available at <https://www.project-syndicate.org/commentary/when-will-lockdowns-be-worse-than-covid19-by-peter-singer-and-michael-plant-2020-04>.
- Steuer B. Aggregation, balancing, and respect for the claims of individuals. *Utilitas* 2021a;33(1): 17–34.
- Steuer B. Contractualism, complaints, and risk. *Journal of Ethics and Social Philosophy* 2021b;19(2):111–47.
- Suikkanen J. *Contractualism*. Cambridge: Cambridge University Press; 2020.
- Tadros V. Localized restricted aggregation. *Oxford Studies in Political Philosophy* 2019;5: 171–204.
- Teather R. The costs and benefits of coronavirus policies need to be weighed. Institute of Economic Affairs blog 2020 Mar 31. Available at <https://iea.org.uk/the-costs-and-benefits-of-coronavirus-policies-need-to-be-weighed/>.
- Tomlin P. On limited aggregation. *Philosophy & Public Affairs* 2017;45:232–60.
- Vaillancourt F. Covid – 19 and the health policy recession: whatever it takes, grandma or the economy or what makes sense? Analysis for Canada. CIRANO Papers 2020pe-01, CIRANO 2020. Available at <https://ideas.repec.org/p/cir/circah/2020pe-01.html>.
- Voorhoeve A. How should we aggregate competing claims? *Ethics* 2014;125(1): 64–87.
- Voorhoeve A. Why one should count only claims with which one can sympathize. *Public Health Ethics* 2017;10(2):148–56.
- Young T. Has the government overreacted to the coronavirus crisis? *The Critic* 2020 Mar 31. Available at <https://thecritic.co.uk/has-the-government-over-reacted-to-the-coronavirus-crisis/>.

---

<sup>1</sup>The first version of this paper was written in May 2020 in the midst of the initial rush of the Covid-19 epidemic. For comments on earlier versions, I'd like to thank Anna Alexandrova, Matti Häyry, Lilian O'Brien, and Fiona Woollard.