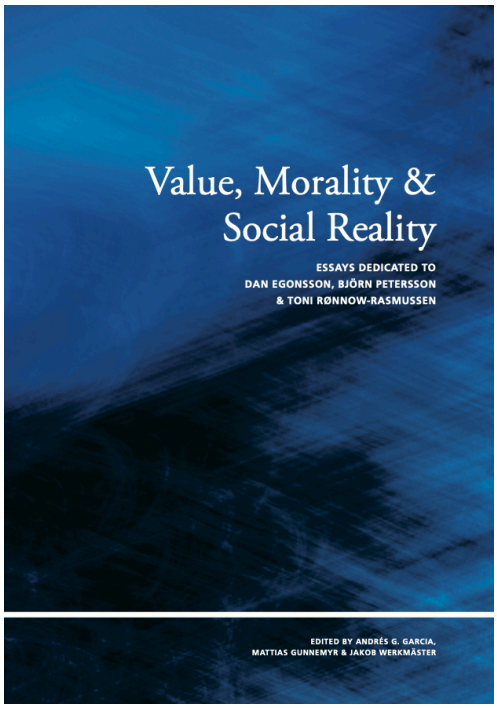
Goodness and Numbers

Wlodek Rabinowicz

In: Garcia, A., Gunnemyr, M. & Werkmäster, J. (2023) *Value, Morality & Social Reality: Essays dedicated to Dan Egonsson, Björn Petersson & Toni Rønnow-Rasmussen.* Lund: Department of Philosophy, Lund University. DOI: https://doi.org/10.37852/oblu.189

ISBN: 978-91-89415-65-2 (print), 978-91-89415-66-9 (digital)



Published by the Department of Philosophy, Lund University.
Edited by: Andrés Garcia, Mattias Gunnemyr, and Jakob Werkmäster
Cover image by Fabian Jones. Cover layout by Gunilla Albertén.

DOI: https://doi.org/10.37852/oblu.189.c533

Goodness and Numbers

Wlodek Rabinowicz[1]

Abstract. You can save either David or Peter and Mary. Is there a compelling reason for saving the greater number? Taurek (1977) (in)famously denied it. In providing such reason one might attempt to establish that it is better if more people survive rather than fewer. This would settle the issue for consequentialists, but even non-consequentialists might find it relevant to the question at hand. The standard worry, however, is that such an axiological claim can only be established by aggregating gains and losses of different persons. As opposed to intrapersonal aggregation, interpersonal aggregation might seem illegitimate. Frances Kamm's Aggregation Argument is meant to overcome this difficulty. I consider how her argument is dealt with by Iwao Hirose, Weyma Lubbe and Rob Lawlor, and what is wrong with it from Taurek's own perspective. But then I suggest that this perspective is untenable: While Taurek correctly analyses the concept of 'better' in terms of fitting preferences, he accunts for fittingness appealing to the wrong kind of reasons. Still, even so, Kamm's argument fails, but a closely related argument may well be acceptable. Unlike the former, that argument recognizes that different ordinary lives typically are on a par; they seldom are equally good.

[1] This essay is an offering to Björn, Dan, and Toni. Toni, I trust, will find it stimulating, as it touches on several value-theoretical themes he has written about. But Björn and Dan might also find it interesting. My primary target is John Taurek, and we have all, I think, found his paper provocative, outrageous, and yet fascinating. I hope these Taurekian qualities are contagious enough to give some borrowed color to my own reflections.

1. Introduction

Consider this choice situation: You can save either David alone, or both Peter and Mary. Those you don't save won't survive. You have no special ties or special obligations to any of those three; they are strangers to you and to each other. Saving Peter and Mary is not more costly than saving David. And, anyway, the costs of saving are modest, let us assume, negligible in comparison with what is at stake. But you can't save all three of them; you must choose.

For definiteness, suppose Peter and Mary are stranded on one desert island, and David on another. The islands are isolated, the food reserves are dwindling, and the stranded individuals have no means of escape on their own. If they stay, they will soon die. You can send a rescue ship to one of the islands, but not to both – not in time to save them all. The islands are too far apart.

Is there a compelling reason for saving the greater number in a situation like this? Should you save Peter and Mary rather than David? Is it what morality requires? In his much-discussed paper, "Should the Numbers Count?" (1977), John Taurek (in)famously denied it. What he would do instead would be to give each person an equal chance of survival, by flipping a coin: If it falls Heads, he saves David; if it falls Tails, he saves Peter and Mary. This would best express "my equal concern and respect for each person" (Taurek, 1977, p. 303)

Actually, Taurek considered a somewhat different set-up: one in which you can save either David or *five* others. And those involved aren't stranded on desert islands. Instead, they are seriously ill and will die unless administered a drug that is at your disposal. One of them, David, needs all of your drug to survive; the other five only need one-fifth of the drug each. These differences between the two choice situations don't matter, I take it, for my discussion. I will continue to focus on the two-islands case, though quotes from Taurek will mention five individuals (instead of just two) whose lives will be lost if David is saved.

To support the claim that you ought to save the greater number, one might argue that it is *better* if more people survive rather than fewer. In the case at hand, this would mean that it is better if Peter and Mary survive than if David alone does. An argument for the former outcome being better would suffice for consequentialists: for them, what you ought to do is to bring about the better outcome. But such an argument would also be of interest to many non-consequentialists. For the latter, deontological considerations, such as the requirement of fairness, might dictate bringing about an outcome that is less than optimal. But many of them might view value differences between alternative outcomes as an important factor that needs to be balanced against other factors that are morally relevant to one's choice. For such non-consequentialists it still is important to determine whether it is better if more people survive, even if this factor need not be decisive. It needs to be weighed against demands of fairness that favor giving each individual an equal chance of continued life.

How then can one argue for it being better if more people survive? A standard worry is that this would require aggregating benefits and/or losses of different persons – that such aggregation is needed for an assessment of the overall value of an outcome. Unlike *intra*personal aggregation, *inter*personal aggregation of benefits and losses might seem problematic. This worry comes to the fore in John Rawls's famous insistence on *the separateness of persons* (Rawls, 1999 [1971], p. 167). Taurek has similar qualms. To begin with, he questions the attempts to aggregate the (purported) objective values of persons or persons' lives to some sort of combined objective value:

> […] when I am moved to rescue human beings from harm […], I cannot bring myself to think of them in just this way. I empathize with [each of] them. […] It is not my way to think of them as each having a certain *objective* value, determined however it is we determine the objective value of things, and then to make some estimate of the combined value of the five as against the one.
>
> (Taurek, 1977, pp. 306f)

He also objects to interpersonal aggregation of personal losses (or gains):

> It is the loss to the individual that matters to me, not the loss of the individual. […] Five individuals each losing his life does not add up to anyone's experiencing a loss five times greater than the loss suffered by any one of the five.
>
> (ibid., p. 307)[2]

Given these worries about the legitimacy of interpersonal aggregation, it is not easy to see how it can be shown that it is better if more people survive rather than fewer. Nevertheless, an attempt to provide such an argument was made by Frances Kamm in the first volume of her *Morality, Mortality*. She called it "The Aggregation Argument." Adjusting the names of the persons involved to our two-islands example, it went like this:

> If (1) it is worse if [Peter] and [Mary] die than if [Peter] alone dies […]; and (2) it is equally bad if [David] alone dies or if [Peter] alone dies […]; then (3) by substitution, it should also be worse if [Peter] and [Mary] die than if [David] alone dies.
>
> (Kamm, 1993, p. 85)

To call it the Aggregation Argument may be somewhat misleading if the argument is meant to establish the conclusion without relying on problematic aggregative premisses. And indeed, in her later work, Kamm began to refer to it as "The Argument for Best Outcomes" (Kamm, 2005, and 2007, pp. 32, 51.). I will, however, continue to use its original name, as it is one under which it has become

[2] Similarly, in his posthumously published reply to Parfit (1978), Taurek questions interpersonal aggregation of pains and sufferings: "pains of […] many cannot be meaningfully summed in a way that has moral significance for preference and choice." (Taurek 2021, p. 315)

widely known, and also because its conclusion (if not its premisses) does appear to involve interpersonal aggregation: the loss suffered by one person is taken to be outweighed by the combined gains of two others.

In this paper, I am going to consider the Aggregation Argument in more detail. I will suggest that while this argument is problematic as it stands, it can be modified to make it defensible. But that modified version, as we shall see, has a somewhat limited reach; it does not extend to all cases in which more people can be saved or fewer.

2. The Aggregation Argument

Kamm's argument was given a more precise form in Hirose (2001) (see also Hirose, 2004, and 2015, ch. 7). What follows is Hirose's reconstruction, with some changes of my own. The argument may be understood as depending on two general principles:

> **Pareto**: An outcome x is better than an outcome y if x is better than y for some individuals and equally as good as y for everyone else.[3]
>
> **Impartiality**: Two outcomes, x and y, are equally good if there is a permutation p on individuals such that, for every individual i, x is equally good for i as y is for $p(i)$.

This formulation of Impartiality differs from Hirose's. In Hirose (201), Impartiality is a principle according to which two outcomes (two "alternatives") are equally good "if they differ only with regard to the identities of the people." And in Hirose (2004), where he refers to Impartiality as "Symmetry", outcomes ("states of affairs") are taken to be equally good if they are transformable into each other by "permutation of personal identities." These formulations work well for outcomes that are fully specified (i.e., for possible worlds) but are problematic if, as in the argument to follow, some of the relevant details are left out from outcome specification. For example, if one outcome is that *I drink wine and you drink beer* while the other is that *you drink wine and I drink beer*, then the former might be better (or worse) than the latter for each of us because of our drinking preferences. Given Pareto, this implies that the two outcomes won't be equally good: the first will be better (worse) than the second. This problem is avoided by the formulation I have suggested. If both for you and for me drinking wine were equally good as drinking beer, then, plausibly, the two outcomes would be equally good. However, relying on my formulation of Impartiality will require adding some supplementary assumptions in the Aggregation Argument.

[3] The name of this principle, "Pareto", might be misleading: In economics, the Pareto condition takes as inputs the outcome preferences of individuals and not how good these outcomes are for the individuals. What an individual prefers may, but need not, coincide with what is good for her.

As here stated, both Pareto and Impartiality ground the impersonal value relations between outcomes – the relations of betterness, period, and of equal goodness, period – in comparisons of how good the outcomes are for different individuals.

In this context, we assume (i) that the compared outcomes involve the same individuals, and (ii) that in these outcome comparisons we bracket non-welfarist considerations. We only consider how good the outcomes are for individuals and either disregard other considerations or take them to be irrelevant for outcome values.[4]

The Aggregation Argument itself has two premisses. The first is implied by Pareto and the second by Impartiality, given some supplementary assumptions. The premisses state impersonal value relations between different outcomes, where an outcome specifies, for each of the three individuals involved, whether that individual survives (+) or dies (-).

Premiss 1: {Peter +, Mary +, David -} is <u>better</u> than {Peter +, Mary -, David -}.

That is, it is better if Peter and Mary survive, while David dies, than if Peter alone survives.

This premiss follows from Pareto, given the supplementary assumptions that it is better for each person to survive, and that a person's survival is at least as good as her death for everyone else, whatever their own fate might be.

Premiss 2: {Peter+, Mary -, David -} is <u>equally as good</u> as {Peter -, Mary -, David +}.

That is, it is equally as good (or bad) that Peter alone survives as that David alone survives.

This premiss follows from Impartiality, given the supplementary assumptions that it is equally good (or bad) for David to be a lone survivor as it is for Peter, and that it is equally bad (or good) for Mary that Peter alone survives as that David alone survives. (Remember that all three individuals are strangers to each other.)[5]

The conclusion of the Aggregation Argument is that it is better if Peter and Mary survive while David dies than if David survives while Peter and Mary die.

Conclusion: {Peter +, Mary +, David -} is <u>better</u> than {Peter -, Mary -, David +}.

---

[4] This welfarist restriction on value comparisons between outcomes will be retained until the last section, where I will very briefly broach the possibility that outcome values might partly depend on considerations of fairness.

[5] The first assumption (as indeed Impartiality itself) presupposes that interpersonal comparisons of how good or bad an outcome is for different individuals are meaningful. But even if we accept this, as I think we should, we might well wonder whether survival has the same value for different individuals. Doesn't this depend on how good their continued lives would be for them? I will return to this issue later, in the final section.

To derive Conclusion from Premisses 1 and 2, we only need to assume that

> **Betterness is transitive across equal goodness:** For all outcomes *x*, *y*, and *z*, if *x* is better than *y*, and *y* is equally as good as *z*, then *x* is better than *z*.[6]

Suppose we let at least as good as be our primitive relation. Intuitively, we understand it disjunctively: to be at least as good is to be better or equally good. These two disjuncts can then be defined in terms of our primitive relation as, respectively, its asymmetric and symmetric parts:

> *x* is better than *y* iff *x* is at least as good as *y*, but *y* is not at least as good as *x*.
> *x* and *y* are equally good iff *x* is at least as good as *y*, and *y* is at least as good as *x*.

If we now assume that at least as good is a transitive relation, then it follows that both betterness and equal goodness are transitive and that betterness is transitive across equal goodness.[7]

Using suitable supplementary assumptions together with Impartiality and Pareto, the Aggregation Argument can be generalized. We can obtain the conclusion that for any two unequally large disjoint groups of people, it is better if everyone in the larger group survives, while everyone in the smaller group dies, than vice versa. It is better if more people survive rather than fewer.

As suggested above, the conclusion of the Aggregation Argument appears to involve interpersonal aggregation. But what about the argument's premisses? Hirose denies that any of the premisses is "aggregative". The principles on which they are based are not aggregative either:

> Neither Pareto nor Impartiality aggregates the claims of […] different people.
> (Hirose, 2001, p. 341; cf. also Hirose 2004, p. 68)[8]

---

[6] Hirose (2001) doesn't explain how the argument's conclusion follows from its premisses. Probably he thinks it is obvious. Kamm (1993) is a little more explicit: she states that the conclusion follows "by substitution".

[7] Hirose's formulation of the Aggregation Argument differs from Kamm's. In the latter, the outcomes were only partially specified. For example, Kamm's first premiss stated that it is worse if Peter and Mary die than if Peter alone dies; David's fate wasn't mentioned. Also, Hirose's first premiss is not about an additional person dying being worse, but about an additional person surviving being better. But these differences don't matter much. We can re-formulate Kamm's original argument in terms of more fully specified outcomes, with one premiss based on Pareto and the other on Impartiality. Kamm's Pareto-based premiss then states that it is worse if Peter and Mary die, while David survives, than if only Peter dies, while Mary and David survive. The Impartiality-based premiss states that it is equally good (or bad) that Peter dies, while David and Mary survive, as that David dies, while Peter and Mary survive. Since 'worse', like 'better', is transitive across equal goodness (given the transitivity of 'at least as good'), we can draw the conclusion equivalent to Hirose's: It is worse if Peter and Mary die, while David survives, than if David dies, while Peter and Mary survive.

[8] While Hirose denies that Impartiality and Pareto are aggregative principles, he considers them both as necessary for aggregation – as conditions that every aggregative moral theory must satisfy (Hirose,

What he claims isn't obvious. He may be right about Impartiality and the Impartiality-based Premiss 2. If we accept this premiss, it is because we accept that David's survival has the same value as for David as Peter's survival has for Peter, and that David's death has the same disvalue for David as Peter's death has for Peter. Premiss 2 thus doesn't seem to require aggregation of benefits and losses of different persons. But what about the Paretian Premiss 1? According to it, that both Peter and Mary survive is better, period, than that Peter alone survives. This betterness judgment, while eminently plausible, seems to depend on the interpersonal aggregation of the survival benefits that go to Peter and Mary: the survival of both is worth more than the survival of just one of them.

I imagine Hirose could reply as follows. Consider the comparison between the two outcomes in question. Since Peter survives in both, they don't differ as far as he is concerned. But then, in this outcome comparison, we may disregard Peter (just as we disregard David, who dies in both outcomes), and focus on Mary.[9] In one outcome she survives, in the other she dies. The former outcome is thus better for Mary, the only person with regard to which the two outcomes differ, and this is why the former outcome is better, period. If we justify our 'better, period' judgment in this way, we don't seem to rely on interpersonal aggregation.[10]

On this interpretation, then, the Aggregation Argument arrives at what looks like an aggregative conclusion from non-aggregative premisses. But then the conclusion perhaps is not aggregative either, appearances notwithstanding?

3. Lübbe's Critique of the Aggregation Argument

Weyma Lübbe subjected the Aggregation Argument to a scathing critique (Lübbe, 2008; cf. also Lübbe, 2015). This argument compares outcomes in terms of betterness, period, and equal goodness, period. But these notions, Lübbe insists, are not applicable to outcomes ("states of affairs") at all. As opposed to betterness or equal goodness *for* a person, there are no relations of betterness/equal goodness,

---

2015, ch. 2). This necessity claim is criticized by Gustafsson (2017), who provides examples of intuitively aggregative theories that violate Impartiality and Pareto, respectively.

[9] Indeed, when arguing that the Impartiality-based Premiss 2 is not aggregative, we have already done something similar: we have disregarded Mary, whose fate is the same in the two outcomes that this premiss compares.

[10] This non-aggregative way of arguing for Premiss 1 can be extended to outcomes in which more than one additional person benefits. We can proceed in steps, each time adding a benefit to yet another person. Since by the argument above, each step is an improvement, the transitivity of betterness implies that the last outcome in this sequence is better than the first.

As Iwao Hirose has pointed out (private communication), the two principles that underlie the premisses (Pareto and Impartiality) are both derivable from Leximin, which intuitively is not an aggregative principle.

period, between states of affairs. Likewise, as opposed to goodness-for, there is no property of goodness, period, that accrues to states of affairs. On her view, judgments of betterness, period, or goodness, period, are essentially *moral* in nature and thus only apply to what can be morally evaluated. In particular, instead of states of affairs, they may apply to *choices* that bring these states about.

> When we go beyond 'better for' judgments [to judgments of betterness, period], we do not in fact evaluate states of affairs but choices that bring about states of affairs, the choices of a hypothetical decision maker. We make, more precisely, moral evaluations. States of affairs […] are not the proper objects of moral judgments. They are not to blame, even if they are very bad for the people involved.
>
> (Lübbe, 2008, p. 74)

Lübbe takes Taurek to be her ally in this. A "Taurekian" will be able to say that a state of affairs is better than another *for a person*, but not, Lübbe claims, that it is better, period.

> […] an impartial observer, a Taurekian one […] would not, of course, answer that [a better choice] brings about a better state of affairs. […] 'better, period' judgments evaluate choices, not states of affairs […]
>
> (Lübbe, 2008, p. 75)

It is debatable whether she is right in her reading of Taurek. The next section will consider this issue in more detail.

So, how does the Aggregation Argument fare if betterness and equal goodness, period, can be predicated of choices, but not of states of affairs? To answer this question, the argument needs to be re-formulated to make choices the objects of evaluations:

> **Premiss 1*:** To choose that Peter and Mary survive (while David dies) is <u>better</u> than to choose that Peter alone survives.
>
> **Premiss 2*:** To choose that Peter alone survives is <u>equally as good</u> as to choose that David alone survives.
>
> Therefore, because betterness is transitive across equal goodness,
>
> **Conclusion*:** To choose that Peter and Mary survive (while David dies) is <u>better</u> than to choose that David alone survives.

What should we say about this 'choice version' of the Aggregation Argument? Lübbe considers Premiss 1* to be highly plausible, and she thinks Taurek would concur. But what about Premiss 2*? Evaluative comparisons of choices may well depend on what other choices are at the agent's disposal. This has repercussions for Premiss 2*. A moment's reflection suffices to realize that this premiss is glaringly

false. It would be outrageous to choose to save Peter alone, letting Mary die, when we can save her along with Peter. But there is no such outrageous omission if we choose to save David alone. There is no one whom we can save along with David.

> […] in choosing Y [= the survival of Peter alone] we decide deliberately to watch Mary die and waste a resource that could have been used to save her, while in choosing X [= the survival of David alone] we do no such thing […] For Y there is a Pareto improvement open to choice […]. For X there is none.
>
> (Lübbe 2008, p. 80)

Thus, to choose that Peter alone survives is not equally as good as to choose that David alone survives. It is *worse*. But then, in this choice version, the argument crumbles. While choosing that both Peter and Mary survive (z) is better than choosing that only Peter survives (y), choosing that only David survives (x) is also better than y. Therefore, we cannot reach the conclusion that choice z is better than choice y.

Lübbe writes:

> I conclude that [… a] Taurekian will deny the second premise. Note that he would not deny it if X and Y were the only open choices. (ibid.)[11]

At this point, one might demur: Isn't it rather natural to interpret choice comparisons as implicitly presupposing that the hypothetical decision maker must make one of these choices? That they are the only ones that are open?

I am not sure that we do presuppose it, but suppose we do. Then Premiss 1* would still be plausible: if we can either save Peter and Mary or just Peter, choosing to save both would be better. And Premiss 2* would now also be plausible: if we can only save one person, Peter or David, Lübbe believes that either choice would be equally good.

Would this suggestion save the choice version of the Aggregation Argument? It would not; its conclusion would no longer follow from the premisses. The reason is that, on this interpretation, the two premisses refer to different choice situations. When applied to choices, betterness must be transitive across equal goodness only if the choice situation is held *constant*. If z is a better choice than y in one choice situation, and x is equally as good as y in another, in which z is unavailable, then there is nothing to guarantee that z is better than x when both x and z are open to the agent.

[11] But what if another person, Ellen, were stranded on the same island as David and we could save her along with David? Then Premiss 2* would become plausible. Choosing that only David survives would be equally outrageous, equally as bad, as choosing that only Peter survives. Lübbe should be willing to accept the choice version of the Aggregation Argument for this case and conclude that it would then be better to choose to save the greater number: better to save Peter and Mary than David alone.

Thus, this attempt to save the choice version of the Aggregation Argument fails. Lübbe is right; if judgments of betterness and equal goodness, period, are applicable to choices but not to the states of affairs these choices bring about, the Aggregation Argument doesn't go through; either one of its premisses is false, or its conclusion doesn't follow from the premisses.

4. Taurek's view

But did Taurek hold the view ascribed to him by Lübbe? Did he think that judgments of betterness, worseness, or equal goodness, period, do not apply to states of affairs?

It might seem that he did. Consider the following quotes (remember that Taurek considers a situation in which we can save either David or five other persons):

> The claim that one ought to save the many instead of the few was made to rest on the claim that, other things being equal, it is a worse thing that these five persons should die than that this one should. It is this evaluative judgement that I cannot accept. […] I do not wish to say this unless I am prepared to qualify it by explaining to whom or for whom or relative to what purpose it is or would be a worse thing.
>
> (Taurek, 1977, pp. 303f)[12]

> Some will be impatient with all this. […] They will insist that I say what would be a worse (or a better) thing, period. It seems obvious to them that from the moral point of view, since there is nothing special about any of these six persons, it is a worse thing that these five should die while this one continues to live than for this one to die while these five continue to live. It is a worse thing, not necessarily for anyone in particular, or relative to anyone's particular ends, but just a worse thing in itself.
>
> (Taurek, 1977, p. 304)

To this Taurek responds:

> I cannot give a satisfactory account of the meaning of judgments of this kind. (ibid.)

In his paper on Taurek, Rob Lawlor suggests that Taurek for this reason would also "simply deny" the Pareto-based premiss of Kamm's Aggregation Argument: "he will deny that [Peter] and [Mary] dying is worse than [Peter] alone dying." (Lawlor, 2006, p. 152)

---

[12] Christian Piller, who is sympathetic to Taurek's views, puts these words in David's mouth (paraphrasing Taurek, 1977, p. 299): "What do you mean when you say that it would be worse if the many died than if I died? It would be *worse for me* if you saved them and it would be *worse for each of them* if you saved me. That's all there is to it." (Piller, 2014, p. 184) And Piller continues: "If goodness-for points in different directions, there is no resolution of this conflict by any simple appeal to goodness." (ibid.)

Lawlor explains how he thinks Taurek would deal with Pareto-improvements, without falling back on 'better, period' or 'worse, period' judgments – the kind of judgments that "Taurek cannot make sense of" (Lawlor, ibid.):

> […] if it is a choice between [Peter] alone dying or [Peter] and [Mary] both dying, then Taurek will agree that the latter outcome is worse for [Mary], but he will deny that it is worse, period. Taurek will agree that we should save [Mary], even if we can't save [Peter]. But – for Taurek – this is not because two dying is worse than one dying. Rather, we should save [Mary] simply because it is better for [Mary] if [she] lives.
>
> (Lawlor, ibid.)[13]

But do Lübbe and Lawlor interpret Taurek correctly? Does Taurek really question the meaningfulness of 'better, period'-judgments regarding outcomes? Or does he only question whether the judgments such as his *opponents* want him to accept can be given a meaning that would make them plausible?

What strongly supports this second interpretation is that, immediately after declaring that he "cannot give a satisfactory account of the meaning of judgments of this kind," Taurek proceeds, in the same paragraph, to give a general account of what it involves to judge that one outcome is worse, or better, period, than another and how it differs from judging it to be better for a person or a group:

> When I judge of two possible outcomes that the one would be worse (or better) for this person or this group, I do not, typically, thereby express a preference between these outcomes. Typically, I do not feel constrained to admit that I or anyone *should* prefer the one outcome to the other.[14] But when I evaluate outcomes from an impersonal perspective (perhaps we may say from a moral perspective), matters are importantly different. When I judge that it would be a worse thing, period, were this to happen than were that to happen, then I do, typically, thereby express a preference between these outcomes. Moreover, at the very least, I feel constrained to admit that

[13] Lawlor here refers to Kamm's original formulation of the Aggregation Argument. Undoubtedly, he would say the same about the Paretian Premiss 1 in Hirose's version. On Lawlor's view, Taurek would deny that Peter and Mary surviving is better than Peter alone surviving. Rather, we should also save Mary if we save Peter simply because it is better for Mary if she lives.

[14] Does this mean that, for Taurek, 'better for' judgments are purely descriptive – devoid of any normative force? Possibly. Alternatively, he might treat them as conditionally normative, on the lines later developed by Stephen Darwall. Darwall suggests that what is good or better for a person is what one ought to want for her *if* one cares for her. (Cf. Darwall, 2002, pp. 4, 8, 6-7, 8-9, 26 and 48.) This account may be contrasted with a more categorically normative analysis proposed by Toni Rønnow-Rasmussen. On that analysis, *x* is good for a person *i* iff one ought to favor *x* for *i*'s sake. This normative requirement is not conditioned on one's concern for *i*. (Cf. Rønnow-Rasmussen, 2004, 2007, 2011.) Unlike Darwall's, Rönnow-Rasmussen's proposal is definitely in conflict with Taurek's view.

> I *should* have such a preference, even if I do not. It is a moral shortcoming not to prefer what is admittedly in itself a better thing to what is in itself a worse thing.
> (Taurek, 1977, pp. 304f)[15]

If one reads this passage as an analysis (complete or at least partial) of judgments that one outcome is better, period, than another, it suggests the following hybrid account, with an expressivist and a cognitivist part:

> A judgment that an outcome *x* is better, period, than an outcome *y*
>
> (i) typically *expresses* a preference for *x* over *y*;
>
> and
>
> (ii) *states or implies* that one ought to prefer *x* to *y*.

In clause (ii), "one" is to be understood as "everyone", even though Taurek in the quoted passage only considers what he who makes the judgment should prefer. But he considers it a moral shortcoming not to have this preference with respect to what is "in itself a better thing". Which suggests that it is a preference that everyone should have. This reading is fully confirmed by Taurek's subsequent discussion, which I will consider in section 6.[16]

5. Fitting-Attitudes Analysis

The cognitivist part in Taurek's account of 'better, period' judgments (clause (ii) above) is in line with the *fitting-attitudes analysis of value* (*FA-analysis*, for short),

---

[15] Lawlor, who quotes the same passage, feels compelled to admit that "[i]f we interpret Taurek in this way, he can no longer resist the claim that five dying is worse than one dying by simply insisting that such statements don't make sense." (Lawlor, 2006, p. 304)

See also the following passage from Taurek's posthumously published reply to Parfit (1978):

> For example, I really do think it would be, morally speaking, a better thing if this one person were to suffer some substantial pain if these many others could each thereby be spared an agonizing pain. The one alternative is to be preferred to the other. […] It doesn't matter what role I imagine myself to occupy in this situation. (Taurek, 2021, p. 319)

[16] Cf. also these quotes from Taurek (2021), pp. 319 and 320, respectively:

> "In 'Should the Numbers Count?' I traded on what for me is an inextricable connection between the thought of one thing's being, from a moral point of view, preferable to its alternative, and the thought that one should prefer it. Anyone should, for it is, morally speaking, preferable."
>
> "[…] when one alternative, seen from this impersonal perspective, is judged morally preferable to another, then that is what we should prefer; that's what anyone contemplating these same alternatives should prefer."

according to which to be valuable is to be a fitting target of a pro-attitude. FA-analysis has an attitudinal component (pro-attitude) and a normative component. "Fitting" is the standard term used for the latter, but, in some versions of the analysis, it is replaced by a more generic normative expression, such as "ought", "should", or "has reasons to".[17]

For betterness, the pro-attitude that typically figures in the analysis is preference. This suggestion goes back to Brentano, one of the founding fathers of FA-analysis (see Brentano, 1969 [1889], p. 26). Thus, using "ought" for the normative component, we get:

> x is *better*, period, than y iff one ought to prefer x to y.[18]

Equal goodness is accounted for correspondingly, in terms of indifference (equi-preference):

> x is *equally as good*, period, as y iff one ought to be indifferent between x and y.

In Rabinowicz (2008), I developed an FA-modeling and taxonomy of binary value relations (see also Rabinowicz. 2012). Exploiting the idea that, for the normative component of FA-analysis, two levels of normativity are available – strong (ought, required, fitting) and weak (may, permissible, not unfitting) – I showed that there are several types of binary value relations that fall outside the standard trichotomy of better, worse, and equally good. Here is the definition of one such type of relation, *parity*:

> x and y are *on a par* iff one may prefer x to y, and one may prefer y to x.[19]

In other words, preferring x to y is permissible, and so is the opposite preference. The notion of parity will prove helpful in the final section.

Cases of parity typically arise in multi-dimensional comparisons. One item, x, might be superior to the other, y, on some dimensions and inferior on others. To reach an overall comparative assessment of such items, the relevant dimensions need to be weighed against each other. If there is some latitude in such weighing, x might come higher than y in the overall preferential assessment given one admissible

[17] For a short history of FA-analysis, see Dancy (2000) and Rabinowicz & Rønnow-Rasmussen (2004).

[18] Does Taurek accept this equivalence? It is not clear (and, even so, he might not accept it as an account of the *meaning* of "better"). In any case, as we have seen in the preceding section, he does accept that it holds from left to right. This will suffice for the discussion to follow.

[19] For this definition, see Rabinowicz (2008), but the idea of parity as an independent value relation is due to Ruth Chang (see Chang, 2002a, 2002b, 2005).

assignment of weights to dimensions and lower given another.[20] We will then have a case of parity: it is permissible, all things considered, to prefer *x* to *y*, but it also is permissible, all things considered, to prefer *y* to *x*.[21] It will then typically also be permissible to be indifferent between *x* and *y*, although I haven't included it in my definition of parity.[22]

Even if I were to include it, it might well be questioned whether my definition conforms to the ordinary usage of "on a par". Hugh Barrett has recently suggested modifications to my definitions of betterness and parity (Barrett, 2022, ch. 1).

To begin with, he suggests a weakening of the definition of betterness. For *x* to be better than *y* it is enough, in his view, if one may prefer *x* to *y* and ought to prefer it or be indifferent. To put it differently, *x* is better than *y* iff one ought to favor *x* at least as much as *y*, and may favor it more than *y*. Unlike mine, this definition allows for cases in which it is permissible to be indifferent between a better item and one that is worse. While weaker than mine, Barrett's definition still implies, as it should, that betterness is an asymmetric relation.

Secondly, he proposes a very weak definition of parity: two items are on a par iff one may be indifferent between them. This definition makes parity a kind of 'rough equality'. Two items that are on a par might be equally good, or one of them might be (somewhat) better than the other. (Remember that on Barrett's weak definition of betterness, indifference between a better and a worse item might be permissible.) As I have defined parity, none of this is possible. On my account, if two items are on a par, then none of them is better than the other, nor are they equally good. Barrett suggests that what I and others (especially Ruth Chang) have been after is not "on a par", but "merely on a par" (i.e., on a par, but neither equally good nor better or worse). He may be right, but for the discussion that follows it doesn't matter whether the way I use "on a par" conforms to the ordinary usage. What does matter is that it is clear what kind of value relation I have in mind.[23]

---

[20] An overall assessment of an item need not be literally a weighted average of its scores on different dimensions. Still, it is a function (whose parameters admit of some latitude) of how well the item does on each dimension.

[21] It need not be like this in all cases of multi-dimensional comparisons. In some, one item might be ranked higher than the other on every admissible assignment of weights to dimensions. (Either because it is superior on each dimension or because it is inadmissible to give much weight to the dimensions on which it is inferior.) Then preferring it would be required. Which would imply that it is better than the other item.

[22] For all I know, there might exist cases in which it is permissible to prefer one item to the other and permissible to have the opposite preference, but impermissible to be indifferent. Still, even if such cases might exist, they certainly aren't typical.

[23] In what follows I will also allow for parity in personal values. How such parity should be understood depends on how we decide to analyze personal value (goodness-for) in the first place. If we find Darwall's account of 'good for' attractive, we could say that the value of *x* for *i* is on a par with the value of *y* for *j* iff those who care for *i* and *j* may, for *i* and *j*'s sake, prefer *x* to *y*, but they may also prefer *y* to *x*.

How can an FA-account of value relations guarantee that betterness and equal goodness are transitive and that betterness is transitive across equal goodness? These transitivity properties of value relations will follow if we impose transitivity as a rationality constraint on permissible preferences and indifferences (cf. Rabinowicz, 2008), or – alternatively – if we re-interpret the concept of preferences in such a way that transitivity will fall out as their conceptually necessary feature (cf. Rabinowicz, 2012). For details, I refer the reader to these two papers.

6. Back to Taurek

If, contrary to Lübbe's suggestion, 'better, period' judgments regarding outcomes are meaningful on Taurek's view, then why does he deny that it is better, period, that more people survive rather than fewer? Why does he deny that it is better, period, if David dies but five others (or two others, or fifty others) survive than if David survives but the others die?

He denies this 'better, period' judgment because he accepts the FA-connection between what is better, period, and what everyone ought to prefer, but denies that everyone ought to prefer that more people survive rather than fewer. In particular, it is permissible for *David* to have the opposite preference:

> I do not think [David] morally deficient in any way because he prefers the outcome in which he survives and the others die to the outcome in which they survive and he dies.
> (Taurek 1977, p. 305)[24]

The same goes for anyone who has close ties to David:

> In a situation where the one person, David, is a friend of mine and the others strangers to me, I do have a preference for the one outcome as against the other, to me a natural and acceptable preference. […] His survival is more important to me than theirs. I would expect them to understand this, provided they were members of a moral community acceptable to me, just as I would were our roles reversed.
> (ibid.)

Let us apply this to the two-islands case. If preferring David's survival, even if it means that Peter and Mary will die, is permissible for some of us (for David and his friends), then it follows that it is *not* the case that

> (every)one ought to prefer {Peter +, Mary +, David -} to {Peter -, Mary -, David +}.

[24] Cf. also Taurek (2021, p. 320: "That [David] prefers the alternative in which he survives manifests to me no moral deficiency in him. I cannot expect that he should prefer the alternative in which they survive and he dies. Hence I cannot give as my reason for sparing the five instead of him the impersonal evaluational comparison that the one outcome is, from a moral point of view, preferable to the other."

But then, given FA-analysis, it is *not* the case that

{Peter +, Mary +, David -} is better than {Peter -, Mary -, David +}.

Thus, Taurek would reject the Conclusion of the Aggregation Argument. He can do it because he would reject the Impartiality-based Premiss 2:

{Peter +, Mary -, David -} is equally as good as {Peter -, Mary -, David +}.

His objection here would be predictable: It is permissible for David (and David's friends) to prefer that David alone survives rather than that Peter alone survives. Consequently, it is not the case that everyone ought to be indifferent between these two outcomes. Therefore, on the FA-account, it is not the case that these two outcomes are equally good. Premiss 2 is false.

The falsity of one of the premisses suffices for the Aggregation Argument to crumble. But it might still be of interest to consider what Taurek would want to say about the argument's other premiss, Premiss 1:

{Peter +, Mary +, David -} is better than {Peter +, Mary -, David -}.

I am not sure what he would say about this Pareto-based claim. He might say that for people for whom Mary is a stranger, it is permissible to be indifferent between Peter and Mary surviving and Peter alone surviving. This would imply, on his FA-account of betterness, that the former outcome is not better than the latter.[25]

But Taurek might instead say that it would be a moral shortcoming to be indifferent to Mary's plight even if she is a stranger. Pro tanto, one should prefer outcomes that are better for others, especially if they are so dramatically better for them as when it is a matter of life and death. Arguably, this is required by the fundamental respect we owe to other persons.

Would this suffice to guarantee the truth of Premiss 1? What about the *enemies* of Mary? Is it impermissible also for them not to prefer that she survives along with Peter? Would it be a moral shortcoming on their part not to have this preference? As long as this is unclear, Taurek's assessment of the Pareto-based Premiss 1 also remains unclear.[26]

In this section, I have sketched what I take to be Taurek's view regarding the Aggregation Argument. Now let me consider how we can resist it.

---

[25]Though, note that on Barrett's weakened definition of betterness, the permissibility of such indifference would not yet falsify Premiss 1. The latter would still be correct if it also were permissible to prefer that Mary survives, along with Peter but impermissible to have the opposite preference.

[26] Kamm (1993, p. 97, fn. 12) reports that in a conversation she had with Taurek the latter said he would go so far as to accept Pareto. But to derive Premiss 1 from Pareto we had to assume that there is no one for whom it is better that Mary dies.

7. Reasons of the Wrong Kind

In Rabinowicz & Rønnow-Rasmussen (2004), we posed and (unsuccessfully) attempted to solve a worrisome problem for FA-analysis: Some of the reasons for a pro-attitude (or a con-attitude) towards an object might have no bearing on that object's value (disvalue). Such reasons are of the 'wrong kind' as far as the FA-analysis is concerned: their presence doesn't make the object valuable (disvaluable). Here is a dramatic example: A powerful demon demands that we admire him; if we don't, he will destroy the world. His determination (and capacity) to destroy the world unless we comply is a very strong reason for admiring him, but it does *not* make him admirable. Thus, it may be that one ought to have a pro-attitude towards an object (in this case, the demon) even though the object itself lacks value.

Some philosophers have argued that examples like this don't pose a serious obstacle for FA-analysis. One line of resistance is to deny that we do have a reason to admire the demon. What we have is a reason to *desire* that we admire him and reasons to bring it about.[27] This only shows that admiring the demon has value, which is uncontroversial given his threat. Another line of resistance is different: Even if we might have a reason to admire the demon, to admire such an evil creature would not be *fitting*. And on the FA-analysis, an object x is valuable only insofar it is fitting to have a pro-attitude towards x. It isn't enough, on this interpretation of the FA-account, that one ought to have a pro-attitude towards x for x to be valuable. Fittingness is taken to be a distinctive deontic concept that differs from a mere Ought.[28]

While each of these two maneuvers would help to disarm the demon example, I doubt whether they could get FA-analysis entirely off the hook. Unless an attitude's fittingness is understood as its adequacy to the object's value, which would make the analysis of value in terms of fittingness viciously circular, it is not difficult to provide examples of cases in which it intuitively *is* fitting to have a pro-attitude towards an object that is not valuable or a con-attitude towards an object that is not disvaluable. In such cases, it also seems perfectly intuitive to say that we have reasons for these attitudes and not merely reasons for desiring to have them (and for bringing them about. Thus, think of a rather indifferent poem composed by your teenage daughter. It may well be fitting for you as a parent to admire it; not just to pretend admiration. You would be a worse parent otherwise. And the reason to admire this poem is that it is your daughter's creation. Or think of Bernard Williams's example of a lorry driver who runs over a child, despite taking all sorts of precautions. It is fitting for the driver to feel guilt, even though he knows he isn't blameworthy; the child's death

[27] For this view, see, for example, Gibbard (1990), p.37, Parfit (2001, 2011, Appendix A), Skorupski (2007, 2010).

[28] For such a "fittingness first" approach to FA-analysis, see McHugh & Way (2016, 2022) and Howard (2019).

wasn't his fault. Still, there would be something morally wrong with him if he didn't feel guilt and remorse. There is a reason for him to feel guilt: he caused the child's death, however inadvertently. (Cf. Williams, 1981.)

One might put this worry about the effectiveness of the appeal to fittingness as follows: In some cases, what makes attitudes fitting are not the value-making features of their *objects*, but the deontological constraints on the *subjects* of the attitudes: constraints that require the subjects to have these attitudes. From the point of view of FA-analysis, such deontological constraints give rise to reasons of the wrong kind.[29, 30]

The difficulty I have just sketched is a problem for the FA-analysis of value properties, but a similar problem arises for value relations. Thus, it might be that an item *x* is better than another item *y* but we still ought to prefer *y* to *x*: we might have strong reasons for doing so and it might be fitting. Thus, I ought to prefer my daughter's poem to a better poem of her classmate. As a father, I have a reason to prefer it, and it may well be fitting.

In other words, some reasons for preference might be of the wrong kind from the point of view of FA-analysis. They are the kind of reasons that this analysis needs to bracket – exclude from consideration. Which doesn't hinder, of course, that they might be perfectly good reasons, as such, and that the attitudes grounded in such reasons might be required and fitting. But such reasons don't make the preferred outcome better or the dispreferred outcome worse.

In what follows, I refer to them as "WK-reasons" (short for "reasons of the wrong kind"). Since such reasons must be bracketed in the FA-account, we need to add an appropriate proviso to the analysis. As applied to the value relations, this analysis should be framed roughly along the following lines:

> *x* is <u>better</u> than *y* iff, with WK-reasons bracketed, there are conclusive reasons for preferring *x* to *y*.
>
> *x* and *y* are <u>equally good</u> iff, with WK-reasons bracketed, there are conclusive reasons for being indifferent between *x* and *y*.

---

[29] Cf. Crisp (2008), pp. 260f. Crisp emphasizes the role of deontological constraints in generating some of the reasons of the wrong kind.

[30] This problem might not arise if one instead of fittingness appeals to the concept of c*orrectness*. Fitting attitudes might still be incorrect if what makes them fitting are deontological constraints. (Think of my admiration for my daughter's indifferent poem, or of the lorry driver's feelings of guilt.) And indeed, Brentano's original formulation of FA-analysis was framed in terms of correct ("richtig") pro-attitudes (Brentano, 1969 [1889], p. 18). His approach has been recently revived in Danielsson & Olson (2007). The worry is, though, that correctness does better than fittingness only because it is so natural to understand the correctness of a pro- or con- attitude as its adequacy to the object's value. As this would make the correctness version of FA-analysis circular, Danielsson and Olson declare that correctness is a primitive, unanalyzable concept. How plausible is this? To seriously assess their proposal would, however, take us too far from the main topic of this paper.

The 'wrong kind of reasons' problem (the WKR problem, for short) is the problem of defining WK-reasons, in a non-circular way. Obviously, if they need to be explicitly excluded from consideration in the FA-analysis of value, it wouldn't do to define them as reasons for an attitude that do not bear on the value of its object. To provide a satisfactory definition of WK-reasons has proven to be difficult. (Cf. Rabinowicz & Rønnow-Rasmussen, 2004.) But certain types of WK-reasons are easier to identify. When arguing that the death of five persons is not worse than the death of one, David, Taurek appeals to David's *personal* reasons for preferring his own survival to the survival of others: reasons of self-interest. Likewise, he appeals to the personal reasons that David's friend would have for the same preference: reasons of friendship. Such personal, "agent-relative" reasons involve an essential reference to the person who has the reason in question: to her interests, projects, social and institutional roles, personal ties and attachments. (Cf Nagel 1970, 1986.)[31, 32]

It is arguable that all agent-relative reasons are WK from the point of view of FA-analysis of *impersonal* value (goodness, period, and betterness, period). Such an analysis must focus on reasons that are there for *everyone*, and not merely for this or that person. Thus, agent-relative reasons should be bracketed in the analysis.

This allows us to resist Taurek's argument. As we have just seen, in defense of the claim that the death of five persons is not worse than the death of one, David, Taurek appeals to the agent-relative reasons on the part of David and David's friend: reasons of self-interest for the former and reasons of friendship for the latter. As I have suggested in the preceding section, he would also appeal to these reasons when confronted with the Impartiality-based Premiss 2 of the Aggregation Argument. But, if I am right, the appeal to agent-relative reasons is illicit in this context. Such reasons must be excluded from consideration when it comes to judgments of betterness and equal goodness, period. With respect to such impersonal evaluations, they are irrelevant.[33, 34]

---

[31] "If a reason can be given a general form which does not include an essential reference to the person who has it, it is an *agent-neutral* reason […] If on the other hand the general form of a reason does include an essential reference to the person who has it, it is an *agent-relative* reason." (Nagel, 1986, pp. 152f) Nagel's way of drawing this distinction is not unproblematic. For some worries, see Rønnow-Rasmussen (2009).

[32] I will continue to talk about "agent-relative" reasons, since this label is so well-established, even though it is slightly misleading when it comes to reasons for attitudes (rather than actions). Perhaps "subject-relative" would be a more adequate characterization.

[33] In the rescue case, these reasons make it *permissible* for David and his friends to prefer David's survival to the survival of others: these reasons override agent-neutral reasons that would otherwise rule out this preference. They differ from the father's reason for admiring his daughter's poor poem in that the latter reason arguably makes his admiration not simply permissible but positively commendable. Still, in both cases, it is a matter of reasons that need to be bracketed when it comes to impersonal evaluations.

[34] Olson (2009) discusses "the Partiality Challenge" to FA-analysis: "there are circumstances in which some agents have reasons to favour or disfavour some object – due to the personal relations in which

It should be obvious, though, that Taurek himself would reject this suggestion. On his view, David's preference for his own survival, even if it means that others must die, is not morally objectionable. And yet it is based on agent-relative reasons. Consequently, this preference cannot be legitimately disregarded when we evaluate outcomes from the impersonal point of view. But Taurek's critics may retort that the moral permissibility of an attitude does not guarantee its relevance for value judgments. They can here take their lead from Parfit (1978) and agree that there is an "agent-relative permission" for David and his friends to have such preference. While this preference is not morally objectionable, it is justified by reasons that don't bear on impersonal value relations.

8. Back to the Aggregation Argument – Final Treatment

If agent-relative reasons don't count when it comes to impersonal evaluations, if Taurek was wrong on this point, does it mean that the Aggregation Argument now is in the clear? One might well think so. Both premisses of the argument appear to be plausible from the impersonal point of view and its conclusion follows from the premisses. Nevertheless, I find this diagnosis much too quick. The Paretian Premiss 1 is indeed very plausible. But Premiss 2, according to which it is equally as good (or bad) if Peter alone survives as if David alone survives, may well be questioned even if we accept the Impartiality principle on which this premiss is supposed to be based. To derive it from Impartiality, we had to assume that Peter's survival is equally good for Peter as David's survival is good for David. But why should it be so? Surely, how good survival is for a person depends on how his or her continued life would unfold. And David's continued life might well be different from Peter's, perhaps better for David than Peter's continued life would be for Peter. Under these circumstances, Premiss 2 would not follow from Impartiality and the Aggregation Argument would not get off the ground.

I expect protests at this point: The kind of situation we have envisaged in the two-islands case is not meant to involve dramatic differences in how the individuals' continued lives would look like. David and Peter are thought of as people whose prospects do not dramatically differ. That one of them would have a wonderful life if he survived while the other's life would be wretched is simply not in the cards.

---

they stand to the object – without this having any bearing on the value of the object." (ibid., p. 365) Olson considers different ways in which this challenge might be met, one of being close to my proposal: agent-relative reasons for attitudes must be disregarded when it comes to impersonal value. His favorite solution is different; it is based on his joint work with Danielsson (Danielsson & Olson, 2007) and their Brentano-inspired analysis of value in terms of correct attitudes. (The idea being that agent-relative reasons don't bear on the correctness of attitudes.) Yet another solution is proposed in Zimmerman (2011). Here, I abstain from discussing these alternative solutions.

What we envisage is a case in which the continued lives of David and Peter would be as ordinary lives use to be.

I agree. Nevertheless, the assumption that their continued lives would be *equally good* for them seems extremely unrealistic. Ordinary lives differ from each other in a multitude of ways. One life has highs and lows that the other lacks, and vice versa. It is superior to the other life in some respects and inferior in others. And, typically, these comparative advantages and disadvantages do not exactly balance off. Rather than being equally good, different ordinary lives are *on a par* when it comes to their personal value. David's continued life would have value for David that is on a par with the value Peter's continued life would have for Peter. If this is right, then it would be reasonable to suppose that from the impersonal point of view we also have a case of parity here.[35]

Premiss 2 should therefore be rejected and replaced by

New Premiss 2: {Peter+, Mary **-**, David **-**} is on a par with {Peter **-**, Mary **-**, David +}.

In other words, Peter' surviving alone is on a par with David's. One may prefer one outcome to the other or have the opposite preference. And, I suppose, one may be indifferent.[36]

What does this replacement in the second premiss imply for the validity of the Aggregation Argument? The short answer is that Conclusion no longer follows. While betterness is transitive across equal goodness, it is not transitive across parity. Thus, while it is better if both Peter and Mary survive than if Peter alone survives, if the latter outcome is on a par with David alone surviving, we cannot conclude that it is better if Peter and Mary survive than if David alone survives.

That betterness isn't transitive across parity is well known. If two items, x and y, are on a par, then slightly improving (or worsening) one of them, say, replacing x with x^+, typically preserves parity. x^+ is better than x but still on a par with y. Here is an example: Consider two holiday trips to very attractive but also very different locations: one to Peru, with its fascinating remains of ancient civilizations, while the other to Galapagos, with its abundant and unique animal life. Intuitively, the two trips are on a par, and, just as intuitively, this parity would still be preserved if one of the trips, say the one to Peru, were slightly improved (maybe extended with an extra day in Cuzco).

---

[35] This move from parity in personal values to parity in impersonal value is plausible in the case we consider, but we shouldn't accept, as a general principle, that any two outcomes must be on a par if they are on a par for every individual for whom they differ. This general principle might seem intuitive, but it is vulnerable to an objection related to the well-known problem of 'opaque sweetening' (originally posed in Hare, 2010; see also Rabinowicz, 2021). Here, I abstain from presenting this objection, but see Nebel (2020).

[36] Here, I interpret "on a par" in accordance with my definition. But the argument that follows would also go through if parity were interpreted as rough equality, as suggested, for example, by Qizilbash (2007) and Barrett (2022).

Does it mean that we are stuck – that we cannot reach the desired conclusion in the Aggregation Argument? No, I don't think so. Note that the examples such as the one with two holiday trips involve *small* improvements: a small improvement of one of one item typically preserves parity. But it is natural to expect that a *large* improvement will not preserve parity; it will make the improved item better than the other item in the pair. If, instead of the original trip to Peru, we were offered an extended trip to both Peru and Patagonia, this would beat Galapagos! It thus seems that

> if x and y are on a par, and x^{++} is much better than x, then x^{++} is better than y.
>
> (Whether x^{++} is much better than y is another matter. Perhaps not.)

This restricted form of transitivity of betterness across parity saves the Aggregation Argument. That both Peter and Mary survive is much better than that only Peter does. The latter is on a par with David alone surviving. But then we can draw the conclusion that it is better (though perhaps not much better) if both Peter and Mary survive than if David alone survives. We are home![37]

Note that this solution can also help in explaining our intuitions regarding some large-scale rescue cases. Consider a two-islands case in which there are thousand people stranded on one island and slightly more than a thousand on the other. In this

---

[37] Mozaffar Qizilbash calls the following feature "the mark of parity":

> "if two states of affairs are on a par, a significant improvement (worsening) of one makes it better (worse) than the other, while small changes in value do not make one better than the other." (Qizilbash, 2005, p. 423. See also Qizilbash, 2007.)

By contrast, Christian Piller questions (in private communication) the validity of this restricted form of transitivity of betterness across parity. I think this issue is complicated. Parity is gradable; it can be closer to or further away from equal goodness. The further two items on a par are from being equally good, the larger improvement it takes for the improved item to become better than the other item in the pair. And vice versa: the closer they are to being equally good, the smaller improvement is needed. (For a suggestion how to interpret such closeness and its degrees, see Hájek and Rabinowicz, 2022. For another interpretation, see Chang, 2016, last section.) In the case at hand, the argument goes through if the continued lives of David and Peter, while being on a par, are relatively close to being equally good: close enough for the addition of Mary's survival to that of Peter to be a sufficiently large improvement.

There's another issue I'd like to mention here. In a paper with Toby Handfield, I have argued that spectrum arguments for various counter-intuitive conclusions can be blocked by positing incommensurability in some intervening steps of the spectrum sequence, but that this incommensurability must then be *persistent*: it must persist when we continually improve the next item in the sequence in the dimension on which it has an advantage over its predecessor. If the kind of incommensurability that needs to be posited is parity, then such persistency would appear to be a counterexample to the view that parity is not preserved if one of the items that are on a par is considerably improved. (Cf. Handfield and Rabinowicz, 2018, Herlitz, 2020, and Rabinowicz, 2022.) However, I am not convinced that this persistency phenomenon is a genuine counterexample to the "mark of parity". It is such a counterexample only if these continual improvements of the next item in the spectrum sequence can eventually add up to a sufficiently large total improvement. But whether they can or cannot is unclear.

case, it is no longer intuitively obvious that we should opt for saving the (slightly) greater number. We might instead consider tossing a coin to decide which group to save. Our intuitions can be explained along the lines sketched above: If there were equally many people on both islands, thousand people on each, the survival of one group would be on a par with the survival of the other. There are in fact a few more people on the other island. The survival of that surplus, along with the thousand others, makes things better. But not much better. When a thousand people survive, adding a few more survivors is, relatively speaking, a small improvement. And we already know that small improvements typically preserve parity. Consequently, we can conclude that in this large-scale case the Aggregation Argument fails: if the groups are large and don't differ much in size, the survival of one group is on a par with the survival of the other. But then we can just as well toss a coin when deciding which group to save.[38]

Indeed, tossing a coin might well be what we ought to do in this large-scale case. There are strong deontological considerations – considerations of fairness – that favor giving each individual an equal chance of survival. Which tossing a coin, when deciding which group to save, will achieve. If we instead simply send the rescue ship to the island on which the slightly larger group is stranded, we treat the individuals on the other island unfairly: we don't even give them a chance to survive. In a small-scale rescue case, in which only one individual, David, is treated unfairly if we decide to save Peter and Mary, this unfairness is outweighed by the significant improvement in the outcome: 2 lives are saved while the expected number of lives saved if we toss a coin is 1,5 (0,5 x 1 + 0,5 x 2). (Cf. Broome, 1998.) But, as Lawlor (2007) points out, in the large-scale case, if we instead of tossing a coin decide to save the slightly larger group, there will be a thousand individuals who are left to die on the other island without being given a chance to survive. They are all treated unfairly. At the same time, the gain in saved lives, as compared with the expected number of lives saved if we instead were to toss a coin, is very limited, possibly not larger than the corresponding gain in the small-scale rescue case. Consequently, in this large-scale rescue case, fairness wins: randomization is required.[39]

An argument similar to Lawlor's was earlier presented by Hirose (2004).[40] But Hirose, unlike Lawlor, takes considerations of fairness to affect the very value of an

[38] A coin toss may be seen as a lottery on outcomes among which it decides. And given my analysis of parity, the expected outcome of a lottery on outcomes that are on a par must itself be on a par with these outcomes.

[39] Lawlor assumes that the relatively small gain in the expected number of saved lives would still make the outcome better in this large-scale case, but that this improvement in the outcome value is outweighed by so many individuals being treated unfairly. Clearly, if the outcomes instead are taken to be on a par, whatever we decide, this fairness argument in favor of randomization is strengthened even further.

[40] He presented it earlier that year, in the doctoral dissertation he defended at St. Andrews. I had the pleasure of being his external examiner.

outcome: an outcome is worse if it involves an unfair treatment. And the more unfairness it involves, the more individuals are unfairly treated, the worse it is. Lawlor, on the other hand, separates the value of an outcome from the assessment of the action that brings it about, with fairness considerations being relevant to the latter but not to the former. So do I. Without this separation, Lübbe's persuasive critique of the choice version of the Aggregation Argument would be applicable to the outcome version as well. For Peter alone to survive is particularly unfair to Mary, given that she could have been saved along with him. Thus, with unfairness counted in, this outcome would be worse (and not merely on a par) than if David alone were to survive. While the conclusion of the Aggregation Argument would still be plausible, the argument itself would be beyond repair.

To sum up: Taurek believed that in rescue cases it is justified to make a decision by a coin toss or some other random process. He was wrong. In small-scale rescue cases, it is better if more people survive rather than fewer. And this has implications for what we ought to do. But Taurek's recommendation seems right in some large-scale rescue cases – cases in which there are many people in each group, all with a claim to fair treatment. If these disjoint groups, only one of which can be saved, don't differ much in size, then denying all the members of the slightly smaller group a chance of survival is not compensated by the relatively small increase in the number of survivors, as compared with the expected number of survivors if the decision is made by a random process. In such large-scale cases, the Aggregation Argument doesn't go through: the outcomes will all be on a par, whatever we decide to do and thus what we ought to do is to toss a coin: considerations of fairness tip the scale in favor of randomization.

Acknowledgements

This paper was long in the making. Its first draft was prepared in 2012, for a colloquium on aggregationism at GAP 8 in Konstanz. The colloquium was organized by Kirsten Meyer and Thomas Schmidt, and the other two speakers were Iwao Hirose and Weyma Lübbe. Then it took ten years before I came round to write the final version. In the meantime, I presented the draft at seminars in Lund, Oxford, Glasgow, York, and Canberra. I am indebted to the participants in these events. Especially, I want to thank my colleagues in Lund, Toni, Björn, Dan, and David Alm, and philosophers in York: Christian Piller, Mozaffar Qizilbash, Thomas Baldwin and Johan Gustafsson. I also wish to thank Iwao Hirose, Weyma Lübbe, Christian Piller, and Thomas Schmidt, who have all sent me challenging and helpful comments.

References

Barrett, Hugh (2022), *Normative Judgement, Rationality, and Reflective Agency*, doctoral thesis (submitted 2021, passed 2022), Canberra: Australian National University, School of Philosophy.

Brentano, Franz (1969 [1889]), *The Origin of Our Knowledge of Right and Wrong*, translated by Roderick Chisholm, London: Routledge & Kegan Paul.

Broome, John 1998), "Kamm on Fairness", *Philosophy and Phenomenological Research* 58: 955-961.

Chang, Ruth (2002a), "The Possibility of Parity", *Ethics* 112: 659-688.

Chang, Ruth (2002b), *Making Comparisons Count*, New York: Routledge.

Chang, Ruth (2005), "Parity, Interval Value, and Choice", *Ethics* 115: 331-350.

Chang, Ruth (2016), "Parity: The Intuitive Case", *Ratio* 29:395-411.

Crisp, Roger (2008), "Goodness and Reasons: Accentuating the Negative", Mind 117: 257-265.

Dancy, Jonathan (2000), "Should we pass the Buck+", in Anthony O'Hear (ed.), *Philosophy, the Good, the True and The Beautiful*, Cambridge: Cambridge University Press.

Danielsson, Sven, and Olson, Jonas (2007), Brentano and the Buck-Passers, *Mind* 116: 511-522.

Darwall, Stephen (3002), *Welfare and Rational Care*, Princeton: Princeton University Press.

Gibbard, Allan (1990), *Wise Choices, Apt Feelings. A Theory of Normative Judgment*, Harvard: Harvard University Press.

Gustafsson, Johan (2017), "Review of Moral Aggregation by Iwao Hirose", *Mind* 126: 964-967.

Hájek, Alan, and Rabinowicz, Wlodek (2022), "Degrees of Commensurability and the Repugnant Conclusion", *Noûs* 56: 897-919.

Handfield, Toby, and Rabinowicz, Wlodek (2018), "Incommensurability and Vagueness in Spectrum Arguments: Options for Saving Transitivity of Betterness", *Philosophical Studies* 175: 2373-2387.

Hare, Caspar (2010), "Take the Sugar", *Analysis* 70: 237-247

Herlitz, Anders (2020), "Spectrum Arguments, Parity and Persistency", *Theoria* 86: 463-481.

Hirose, Iwao (2001), "Saving the Greater Number without Combining Claims", A*nalysis* 61: 341-342.

Hirose, Iwao (2004), "Aggregation and Numbers", *Utilitas* 16: 62-79.

Hirose, Iwao (2015), *Moral Aggregation*, Oxford and New York: Oxford University Press.

Howard, Christopher (2019), "The fundamentality of fit", in R. Shafer-Landau (ed.), *Oxford studies in metaethics*, vol. 14, ch. 10, Oxford: Oxford University Press: 216-236.

Kearns, S., & Star, D. (2008). Reasons: Explanations or evidence

Kamm, Frances M. (1993), *Morality, Mortality*, Volume 1: *Death and Whom to Save From It*, New York: Oxford University Press.

Kamm, Frances M. (2005), "Aggregation and Two Moral Methods", *Utilitas*17: 1-23.

Kamm, Frances M. (2007), *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*, New York: Oxford University Press.

Lawlor, Rob (2006, "Taurek, Numbers, and probabilities", *Ethical Theory and Moral Practice* 9: 149-166.

Lübbe, Weyma (2008), "Taurek's No Worse Claim", *Philosophy and Public Affairs* 36: 69-85.

Lübbe, Weyma (2015), *Nonaggregationismus: Grundlagen der Allokationsethik*, Münster: Mentis.

McHugh, Conor, and Way, Jonathan (2016), "Fittingness First", *Ethics* 126: 575-606.

McHugh, Conor, and Way, Jonathan (2022), *Getting Things Right: Fittingness, Reasons, and Value*, Oxford: Oxford University Press.

Nagel, Thomas (1970), *The Possibility of Altruism*, Princeton, N.J: Oxford University Press.

Nagel, Thomas (1986), *The View from Nowhere*, New York: Oxford University Press.

Nebel, Jacob M. (2020), "A Fixed-population Problem for the Person-affecting Restriction", *Philosophical Studies* 177:2779-2787.

Olson, Jonas (2009), "Fitting Attitude Analyses of value and the Partiality Challenge", *Ethical Theory and the Moral Practice* 12: 365-78.

Parfit, Derek (1978), "Innumerate Ethics", *Philosophy and Public Affairs* 7: 285-301.

Parfit, Derek (2001), "Rationality and Reasons," in *Exploring Practical Philosophy: From Action to Values*, ed. By D. Egonsson, B.Petersson, J. Josefsson, and T. Rønnow-Rasmussen, Aldershot: Ashgate: 17–41.

Parfit, Derek (2011), *On What Matters*, vol. 1, Oxford: Oxford University Press.

Piller, Christian (2014), "What Is Goodness Good For", *Oxford Studies in Normative Ethics*, vol. 4, ed. By Mark Timmons, Oxford: Oxford University Press: 179-209.

Qizilbash, Mozaffar (2005), "The Mere Addition Paradox, Parity and Critical-level utilitarianism", *Social Choice and Welfare* 24: 413-441.

Qizilbash, Mozaffar (2007), "The Mere Addition Paradox, Parity and Vagueness", *Philosophy and Phenomenological Research* 75: 129-151.

Rabinowicz, Wlodek (2008), "Value Relations", *Theoria* 108: 18-49.

Rabinowicz, Wlodek (2012), "Value Relations Revisited", *Economics and Philosophy* 28: 133-164.

Rabinowicz, Wlodek (2021), "Incommensurability Meets Risk", in *Incommensurability: Vagueness, Parity and Other Non-conventional Relations*, ed. by H. Anderson and A. Herlitz, New York and London: Routledge: 201-230.

Rabinowicz, Wlodek (2022), "Can Parfit's Appeal to Incommensurabilities Block the Continuum Argument for the Repugnant Conclusion?", in *Ethics and Existence: The Legacy of Derek Parfit*, ed. by J. McMahan, T. Campbell, J. Goodrich and K. Ramakrishnan, Oxford: Oxford University Press: 430-460.

Rabinowicz, Wlodek, and Rønnow-Rasmussen, Toni (2004), "The Strike of the Demon: On Fitting Pro-attitudes and Value", *Ethics* 114: 391-423.

Rabinowicz, Wlodek, and Rønnow-Rasmussen, Toni (2006), "Buck-Passing and The Right Kind of Reasons", *Philosophical Quarterly* 56: 114-120.

Rawls, John (1999 [1971]), *A Theory of Justice*, revised edition, Cambridge, Mass.: Harvard University Press.

Rønnow-Rasmussen, Toni (2004), "Buck-passing Personal Values", in W. Rabinowicz and T. Rønnow-Rasmussen (eds.), *Patterns of Value – Essays on Formal Axiology and Value Analysis,* vol. 2, Lund: Lund Philosophy Reports 2004:1: PAGES?

Rønnow-Rasmussen, Toni (2009), "Normative Reasons and the Agent-neutral/Relative Dichotomy", *Philosophia* 37: 227-243.

Rønnow-Rasmussen, Toni (2007), "Analysing Personal Value", *The Journal of Ethics* 11: 405-435.

Rønnow-Rasmussen, Toni (2011), *Personal Value*, Oxford: Oxford University Press.

Skorupski, John (2007), "Buck-Passing about Goodness", in *Hommage à Wlodek: Philosophical Essays Dedicated to Wlodek Rabinowicz*, ed. D. Egonsson, J. Josefsson, B. Petersson and T. Rønnow-Rasmussen, www.fil.lu.se/hommageawlodek.

Skorupski, John (2010), *The Domain of Reasons*, Oxford: Oxford University Press.

Taurek, John (1977), "Should the Numbers Count?", *Philosophy and Public Affairs* 6: 293-316.

Taurek, John (2021), "Reply to Parfit's 'Innumerate Ethics'", in *Principles and Persons: The Legacy of Derek Parfit*, ed. By J. McMahan, T. Campbell, J. Goodrich, and K. Ramakrishnan, Oxford: Oxford University Press, ch. 14: 311-322.

Zimmerman, Michael J. (2011), "Partiality and Intrinsic Value." *Mind* 120: 447-483.