

10

Can't Kant Count?

C10

Innumerate Views on Saving the Many over Saving the Few



Sergio Tenenbaum

C10S1

1. The Specter of Numeracy

C10P1 Let us start with a version of a well-known vignette from the ethical literature:

C10P2 **Basic Case**

I am in my boat in a bay which forks into stream A and stream B. I just received a distress signal from two different boats: one carrying two complete strangers (whom I name “Larry” and “Mary” for easy reference) and another carrying one other person, also a complete stranger (whom I name “Jeri”). If I row to stream A, I can save Larry and Mary, but Jeri will die; if I row to stream B, I can save Jeri, but Mary and Larry will die.



C10P3 It is a popular view that in such a situation, I am under obligation to save Larry and Mary. Generalizing the point, we can formulate the following principle:

C10P4 *The More the Merrier*

In a situation in which you can either save a group of n strangers or m strangers at no great effort to yourself, in the absence of any partiality or other morally relevant consideration, you ought to save n strangers if $n > m$.



C10P5 Maximizing consequentialists have an easy time explaining this principle: assuming that that the expected value of saving the life of a person is the same (at least as long as I know no other relevant information about them), the expected value of saving two lives is greater than the expected value of saving one life. In fact, this rationale might seem pretty compelling: it is just

common sense that it is better to save more lives over fewer. Other types of ethical theories have a much more difficult time defending the principle; in fact, some philosophers who reject impartial consequentialism are willing to accept that in **Basic Case**, it is permissible to save Jeri, or that the right action is something different from choosing the outcome in which I save more lives. Famously, John Taurek (1977) argues that there is no obligation to save the many against the few, suggesting that he would flip a coin if he were caught up in this situation.

C10P6 Taurek rejects the central idea that he takes to ground *The More the Merrier*; namely, that there is an impersonal concept of “good *simpliciter*” that applies to outcomes or states-of-affairs, such that the more lives we save, the better the outcome or the state-of-affairs is. Saving Larry’s life is better *for Larry* and saving Mary’s life is better *for Mary*, but there’s no such thing as something being better *simpliciter* or better-for-Larry-and-Mary-and-Jeri, so that we can say that saving Larry and Mary is better than, say, saving Jeri. But this kind of “Geachean” skepticism about “good *simpliciter*” is not the only view that finds it difficult to vindicate *The More the Merrier*. Kantian views committed to the unconditional value of rational agents seem equally inhospitable to the principle, and contractalist views also face difficulties in generating principles that would deliver the same verdict as *The More the Merrier*. All these non-consequentialist views seem committed to the claim that in situations in which our options are to bring about that either n lives are saved or that m lives are saved, the fact that n is greater than m does not, at the most fundamental level, create an obligation to save n lives over m lives; at least, not on the simple grounds that it is better to save more lives. In other words, they are committed to what Parfit calls “innumeracy”.

C10P7 Innumeracy is not strictly incompatible with *The More the Merrier*, since it only makes a claim about what is true at the most fundamental level, but the tension is clear. Taurek is happy to reject principles like *The More the Merrier*, but others have been less cavalier about it; many find it extremely implausible that we should not favor saving the many over the few. However, given the tension between the principle and innumeracy, it is tempting to see non-consequentialist attempts at vindicating it as ad hoc maneuvering to accommodate what are essentially consequentialist intuitions.

C10P8 My aim here is to propose a middle way between Taurek’s bold denial that it is preferable to save the many in our **Basic Case** and endorsing a (seemingly) consequentialist principle.

C10P9 Although, philosophers generally reject Taurek’s conclusion that it is permissible to save the few, I will first present a series of cases in which no

significant partiality is involved, but in which there seems to be nothing wrong with saving the few.¹ If this is correct then we should not grant that *The More the Merrier* is true in full generality. Yet, I think there is a very general argument in favor of accepting that we ought to save Larry and Mary in **Basic Case** that neither depends on, nor vindicates, *The More the Merrier*. In other words, I'll argue that we need not sneak in any form of consequentialism to conclude that we ought to save the many, at least for an important subset of cases. Before we move on, a couple of caveats. Contemporary work on non-consequentialist aggregation has also considered questions of permissible trade-offs in cases in which the harm that I can prevent is different for different people.² These are interesting and important issues that need to be addressed by the non-consequentialist but I cannot do justice to them within the scope of this paper. Similarly, I'll leave aside questions about how to extend any of the results here to contexts of risk and uncertainty.³ Finally, although I will mostly assume a Kantian framework for ease of presentation, the argument is compatible with different versions of non-consequentialism; it should thus be greeted as good news by a large sector of the non-consequentialist population. In fact, the variations on **Basic Case** I'll present should give pause to any non-consequentialist (or anyone at all, I think) who would be willing to endorse *The More the Merrier*. If I am right that this principle leads us astray in such cases, then other non-consequentialists should prefer my proposal for handling **Basic Case** over any view that is committed to *The More the Merrier*.

C10S2

2. Saving the Few

C10P10

Most philosophers agree that if Jeri is your beloved daughter or spouse, you are permitted to save her rather than two strangers.⁴ Similarly, if Larry, Mary, and Jeri find themselves in this bind due to Larry's and Mary's

¹ Taurek says that he would flip a coin if faced with a situation where he could give a lifesaving pill to one person or five persons, but he does not seem to deny that it would also be permissible to save just one person. My argument is that in many circumstances we *are permitted to save the few*, and not just to flip a coin, and that in the cases that Taurek looks at we are *obligated to save the many*. So I'll not be discussing, except *in passim*, questions about the fairness of this lottery as opposed to others. See Saunders (2009) for a discussion of these ideas.

² See Tomlin (2017) and Horton (2018) for skepticism about "limited aggregation" in these scenarios; for a recent positive proposal, see Steuwer (2021).

³ For my own views on this matter, see Tenenbaum (2017).

⁴ But they disagree on what types of partiality permit you to save the few. On Setiya's view (Setiya, 2014), since we do not need a reason to love someone, there is no limit on how wide the

218 SERGIO TENENBAUM

perfidious actions, I might be permitted, or even obligated, to choose to save Jeri. But are situations involving partiality (or involving other independently weighty considerations such as liability) the only ones in which I am permitted to save the few? Let us look at a variation of Peter Singer's (Singer, 1972) famous case of saving a child at the expense of valuable clothing:

C10P11 **Rescuing Baby Chris**

Sam is near a pond and her phone rings. A remotely located lifeguard says: "Our cameras show you near El Pond. This is fortunate; you can rescue Baby Chris! Chris will soon drown unless you come to his aid. I know you can't swim, but the water is not too deep for you, so you can easily walk and fetch little Chris. Just make sure to keep your shoes on, otherwise you won't be able to make it through the rocks to save the toddler." Sam locates Chris, waves at him, he smiles back and she feels very happy she can save this baby. Of course, Sam does not care that she is going to wreck her brand new very expensive shoes; she just goes into the water. Sam starts wading towards Chris when the lifeguard calls her again: "You won't believe this; there are *two* babies drowning in a pond not far from you. You don't have much time left; leave Chris behind and go save the two babies." Sam refuses to leave Chris behind, and continues in her rescue.



C10P12 Had Sam known before she started her rescue mission that there were two babies in an adjacent pond (while only Chris was to be saved in this one), she would certainly, *pace* Taurek, have gone to the adjacent pond. But it now seems permissible that Sam refuses to change course, and insists on rescuing Chris. Sam entered the pond to save Chris and now having settled on this course of action, it seems reasonable that she is not willing to leave him behind. Of course, as a virtuous agent, she must regret that she cannot save the other babies. But she can only do that if she abandons Chris to his death, and she now (reasonably) feels she must ensure that Chris is safe, even if she could save two other babies instead.

C10P13 Two observations are important here. Sam's relationship with Chris is quite minimal. If one accepts that Sam is permitted to continue,⁵ it is not

partiality permission is, so his view "collapses" into Taurek's (But Setiya (2023) defends a view closer to my view here). Parfit (1978) rejects the claim that casual acquaintances can ground these types of permissions.

⁵ Some philosophers with whom I discussed this case think you are *obligated* to save Chris. For my purposes, it is enough to accept a permission.

because of any kind of partiality towards Chris, but simply because Sam *set up to save Chris in particular*. This conclusion seems to fit well with the Kantian view: on the Kantian view, a person has “dignity”, rather than a “price”. At the core of the distinction is the idea that you can “trade up” when things have a price, but not when they have dignity. Suppose we had instead a less virtuous person in the pond’s vicinity (Sammy); Sammy would like to save Chris, but Sammy can’t bear the thought of wrecking such expensive shoes. Suppose now that noticing Sammy’s hesitation, the life-guard offers Sammy identical shoes and a significant sum of extra money. Sammy would probably be irrational if he were to turn down this offer, even ignoring the importance of saving Chris; after all, Sammy would be just trading up. But when Sam is offered the opportunity to rescue two toddlers, she would not be trading up if she accepted the deal; that which has dignity, “cannot be brought into comparison or competition at all, without assaulting its holiness” [4:435]. On the other hand, Sam’s decision to save Chris does not sit well with numerate views; the fact that Sam is on her way to save Chris, the fact that Chris smiled at Sam, or the fact that Sam is thirty feet closer to Chris: none of these properly compensate for the fact that the number of lives to be saved is cut in half.

C10P14

Someone could protest that the judgment that it is permissible to save Chris rests on an illusion; perhaps the inclination to judge it permissible to save Chris is an instance of some kind of “sunk cost fallacy”.⁶ But here lies a dangerous path. Is it essential to the scenario that all these toddlers are threatened by drowning and will be saved by Sam’s personal efforts? Would it be permissible for Sam not to give her expensive shoes to someone else who needs them to barter for a canoe that will allow them to save two other toddlers? It is hard to see how the claim that it is impermissible that Sam save Chris is compatible with our ordinary judgments about the duty of rescue in Singer’s original example. After all, one can always sell the clothing that will be wrecked in the rescue process and save more lives with the proceedings; it might be difficult to give up the idea that Sam can save Chris without radically changing our views about more ordinary rescue cases.⁷

⁶ But note that the “sunk cost fallacy” explanation would predict that we would think that Sam is justified in continuing even before she had any acquaintance with who was being saved in the pond. I think this prediction is false; had she found out that she could save two babies instead halfway to the pond, the situation would be indistinguishable from our *Basic Case*.

⁷ It is quite difficult to delineate precisely the difference between the duty of rescue and other duties of aid (see, for instance, Ignieski (2006)). My point is that we are already committed to reject in other situations that being able to save more lives always trumps the duty to save *this life here and now* even when all involved are strangers.

- C10P15 Of course, none of this is conclusive. However, at the very least, in such cases one cannot accuse the Kantian position of offending against common sense for allowing us to save Chris even if one were inclined (correctly to my mind) to regard it as an embarrassment if it implied the permissibility of saving the few in **Basic Case**. But I think this setup also helps us display the power of the Kantian position. Taurek asks whether it is permissible to save David, my casual acquaintance,⁸ instead of saving more people. From a perspective prior to engaging in any action, it does seem too flimsy a relation to override my reason to save one more person.⁹ However, things look different if I am already engaged in the action of saving the one. On the Kantian view, the reasons that Sam has to save two toddlers do not defeat the reason Sam has to save Chris. The lifeguard has not given Sam a reason that obligates her to stop what she is doing. Of course, the same relations of defeat apply before Sam starts engaging in the action (if, say, Sam knew all along that these two options were open to her). But when Sam is *deciding* between saving one or saving two, some other considerations may apply. Sam might need to act on a principle that is justifiable to everyone involved, or find a reason, or a selection criterion, that is fitting to the momentous decision she faces. Or, as I will argue, there might be something essentially different about the structure of the agent's choice in a case like **Basic Case**.
- C10P16 Of course, this all needs further grounding and explanation. In the next sections I try to explain more precisely how the Kantian can distinguish between **Basic Case** and **Rescuing Baby Chris**.

C10S3 **3. Kantianism, Incomparability, and Choice Principles**

- C10P17 According to Kant, a person has “dignity, that is, an unconditional, incomparable worth” (4:436). In the kingdom of ends:
- C10P18 everything has either a *price* or *dignity*. What has a price can be replaced by something else . . . What . . . is raised above all price and therefore admits of no equivalent has dignity. (4:434)
- C10P19 Of course, I will not attempt here a detailed interpretation of Kant's notion of dignity. But in saying that there is no *equivalent* to the value of a human

⁸ Setiya (2014) helpfully describes David as the barista at my local cafe in order to illustrate the nature of the relation.


⁹ More on this issue later.

being, to their dignity, I think Kant is committed to denying that the value of two human beings is *equal*, at least in an important sense.¹⁰ This is exactly what distinguishes dignity from price: things that have a price can be replaced by things of equal value, but the “unconditional worth” of human beings does not allow for similar trade-offs. Obviously one person has neither *more* nor *less* value than another. But it seems that Kant also wants to rule out that they are of “equal value” if we understand as consequence of this evaluation that one could be replaced by another.

C10P20 It is notoriously difficult to determine what the dignity of each person implies in terms of the permissibility of various actions. But I think we can sketch some basic consequences of Kant’s notion of dignity for our purposes. Suppose I am about to save one person, and you point out that I might save two people instead. “Why would the fact that I can save two people create an obligation to change what I was planning to do?”, I ask. You answer as follows: whatever value there is in saving the life of one person, there will be significantly more value (twice as much?) in saving the lives of two people. Setting aside partiality, or special circumstances, one must choose the option of saving two lives.

C10P21 This kind of reasoning seems to go against Kant’s claims about the special value of human beings, against the idea that they have dignity. Dignity rules out this kind of calculation of value; the above reasoning treats human beings as if they had a price that could be matched and surpassed by the price that the lives of two human beings fetch. Given that human beings have dignity, the reason to save a human life cannot be defeated by the fact that one could save two lives instead. Kant seems committed to the following thesis:

C10P22 *No Defeat Principle*

The reason I have to save one person is not defeated by the (supposed) greater value of saving two lives. 


C10P23 The *No Defeat Principle* explains why the lifeguard’s appeal does not necessarily defeat the reason that Sam must save Chris. And it is worth mentioning that virtually any innumerate view seems committed to the *No Defeat Principle*. Contractualist views rule out adding the justification of two

¹⁰ Of course, Kant does put immense importance in a principle of equality among human beings, especially in his political philosophy. But this is not to say that people have equivalent values in the sense presented here.

independent people in the manner presupposed by simply adding the values of two lives. In refusing to allow a value that is not a value for someone, Taurek's view is also committed to the *No Defeat Principle*. Thus if we can show how commitment to this principle is compatible with providing an adequate rationale for saving the many in **Basic Case** in terms that are clearly acceptable to all these theories, we will have provided a powerful rationale for saving the many that is compatible with a wide array of views that reject (foundational) numeracy.

C10P24 Of course, this is not all that the Kantian view has to say about our relationship to the lives of others. The Kantian view, like most ethical theories, also imposes a duty to aid. For Kant, benevolence is an imperfect duty, and thus it is a duty of wide obligation. It “has in it a latitude for doing more or less, and no specific limits can be assigned to what should be done” (6:393). This latitude is not unconstrained and at least in some contexts the duty to aid imposes obligations to perform particular acts. For the purposes of this paper, I'll just assume that among these more specific duties, we have a duty of easy rescue that I will formulate as follows:

C10P25 *Minimal Duty to Aid*


In situations in which you can directly save the life of a human being at (nearly) no effort or sacrifice to yourself, you have a duty to save their life. 

C10P26 A few caveats about *Minimal Duty to Aid*. First, what counts as “nearly no effort or sacrifice to ourselves” is left open; obviously, the more we allow as “nearly no effort or sacrifice”, the more demanding the duty will be. Secondly, Peter Singer's famous argument (Singer, 1972) seems to show that a principle like *Minimal Duty to Aid* creates a massively demanding duty to aid requiring that relatively affluent people make huge sacrifices for the sake of those in desperate straits elsewhere. One way of blocking this argument is to challenge Singer's (apparent) assumption that we determine what counts as “nearly no effort” by looking at each action separately rather than examining a pattern of repeated actions. Perhaps what counts as “nearly no effort” depends not only on whether it is “nearly no effort or sacrifice” on this occasion but also on whether a pattern of performing such actions at each occasion will require great sacrifices from me.¹¹ My end here

¹¹ For responses to Singer's argument roughly along these lines, see Thomson (2022); Timmerman (2015).

is not to assess such a view, but to preempt some possible objections to the following principle:

C10P27 *Pareto Principle*

If you can save *A* and *B* instead of just *A* at nearly no effort or sacrifice to yourself, then you should save *A* and *B*. 

C10P28 This principle follows from *Minimal Duty to Aid*; in fact, it is just an instance of *Minimal Duty to Aid* applied to the case in which it is already given that you will save one person. But, of course, we need to be careful not to let this consequence necessarily commit ourselves to accepting Singer's argument. So if we block the argument by accepting a version of the above proposal, then we'll need to understand the extra effort or sacrifice not only by its measure on a particular occasion but also by what it demands of us if we act in similar ways in similar occasions. Other ways of blocking the argument will restrict the principle in different ways. Whatever view one favors here does not matter for my argument. The instances of the *Pareto Principle* that will be relevant for our purposes are rather uncontroversial. For instance, I might need just to extend my arm one more time to save the second person. Some of the cases involve literally no extra effort or sacrifice to save the second person; Taurek himself seems to accept some version of the *Pareto Principle*.

C10P29 In **Basic Case**, I have a reason to rescue Jeri given that her life is in danger. But I also have a reason to rescue Larry (alone), a reason to rescue Mary (alone), and a reason to save Jeri, and a reason to save both Mary and Larry. The *Pareto Principle* ensures that my reason to save Mary and Larry defeats any reason I have to save just Larry or just Mary. *No Defeat Principle* ensures, however, that my reason to save Larry and Mary does not defeat my reason to save Jeri or vice-versa. Since we have undefeated reason to act on either option, neither seems impermissible. So it seems that any view committed to *No Defeat Principle*, including the Kantian view, must accept that there is no obligation to save the many.

C10P30 We could preserve innumeracy and yet not be committed to Taurek's view of **Basic Case** if we can argue that the *Pareto Principle* applies to **Basic Case**. But this might appear to be ruled out by the *No Defeat Principle*. After all, the Kantian view on the dignity of persons, at least as we interpreted it, seems to rule out exactly this kind of comparison between one and two lives. More generally, allowing the *Pareto Principle* to be used in this manner

seems indistinguishable from accepting that the value of two lives is greater than the value of one life; it would be just a form of ushering numeracy in through the back door. However, appearances here are misleading. We *can* use the *Pareto Principle* to deliver an obligation to save the many in the relevant cases. Or so I will argue.

4. Sweetening and No Defeat

C10S4

C10P31 So far the combination of *No Defeat Principle* and *Pareto Principle* seem to imply the following:

- C10P32 1. If Sam has the option of saving baby Chris or baby A and baby B, Sam may save Chris (the reason to save Chris is not defeated by the reason to save baby A and baby B) (given *No Defeat Principle*).
- C10P33 2. If Sam has the option of saving Chris or saving Chris and baby A (or baby B), Sam must save both Chris and baby A (or baby B) (given *Pareto Principle*).

C10P34 Of course, there is nothing special about baby Chris or lacking with baby A or baby B. Either of the above claims would be true if we replace the babies that appear in each position. So for instance the following is also true:

- C10P35 3. If Sam has the option of saving baby A or Chris and baby B, Sam may save baby A (the reason to save baby A is not defeated by the reason to save Chris and baby B).

C10P36 These defeat relations are very similar to those familiar from the literature on incommensurability and incomparability. Take, for instance, Ruth Chang's understanding of the "parity" (Chang, 2002) relation; if A and B are on a par, then it is rational to choose A or to choose B even though A and B are not of equal value or indifferent. Options on a par, unlike indifferent ones, resist "sweetening". That is, if option A and option B are on a par, typically, option A is still on a par when we switch option B with a mildly improved version of it. So if going to restaurant *a* and going to restaurant *b* are on a par, they will still be on a par if we make a small improvement to one of the options (by, say, adding an offer for a free *amuse bouche* to restaurant *a* or a \$1-off coupon to restaurant *b*). Of course, not all the defeat relations are the same, and the Kantian position is significantly

different from the claim that the values of the lives of two people are on a par. In the parity cases, typically significant amounts of “sweetening” will change the relation: so, if restaurant *a* offers the meal for free or if they offer me an annual subscription to the restaurant for the price of the meal, it will probably suffice to make the option of going to restaurant *a* better than the option of going to restaurant *b*. But we cannot assume that the same is true for the Kantian view. The Kantian view might conclude that there is nothing that you can add to the option “saving another baby” that will defeat the option “saving Chris”.¹² However, the overlap on defeat relations is enough to learn an important lesson from a variation on the typical cases of choices among items on a par. In particular, looking at how choices between items on a par behave in certain cases of ignorance will help us see how the *Pareto Principle* applies to **Basic Case**.

C1055

5. Sweetening in the Dark

C10P37 As we said above, the resistance to “sweetening” in the case of alternatives that are on a par parallels the structure of the choices among saving lives in Kantian views. Hare (2010) points out that, in the case of parity, some complications arise if we lack knowledge of the options. Let us look at a slightly modified version of Hare’s central example: My house is on fire and I can direct the firefighters to either rescue my rare and unique Fabergé egg or to save my wedding album. I find it difficult to decide between them; neither seems to be more valuable to me than the other, but they are also not of equal value to me. These value relations remain unchanged if the firefighters can also rescue a \$100 dollar bill if, and only if, they rescue the wedding album (though, of course, I must prefer that they save the wedding album and the \$100 dollar bill over saving just the wedding album).

C10P38 Hare considers an interesting variation of this classic type of example. Suppose that I keep these valuables in identical boxes and I can’t remember which box holds which prized possession; they are equally likely to be in

¹² Although this might seem plausible when the numbers are small, many will balk at this consequence when the numbers get large enough (say, if we could save millions of babies instead of baby Chris). On my view, in order to accept this “intuition”, the Kantian needs to show that at some point the difference in numbers makes a qualitative difference; or, that if the numbers are high enough that the nature of the available actions change. A proper examination of these cases, though, is beyond the scope of this paper (for a related approach for cases involving risk, see my (Tenenbaum, 2017)).

either box. The firefighters can see that there is a \$100 bill next to one of the boxes, and they ask me if I would rather that they save this box. So my options are roughly as follows:

C10T1 **Table 1:** Sweetening in the Dark

Box A	Box B
.5 (album, egg)	.5 (album, egg); \$100

C10P39 In such a case it seems rational for me to ask the firefighters to save Box B. After all, it seems that Box B prospects are superior no matter how the album and the egg compare to each other. But something seems strange here: no matter what is in Box A, if I knew what was in Box A, I would be rationally permitted to choose Box A. Hare imagines a better-informed self who knows what is in each box. I *know* that my better-informed self is permitted to choose it. But then why am *I* not permitted to choose Box A? Hare himself thinks that there are two plausible decision theories at conflict here: prospectism and deferentialism; roughly, a clash between a principle that tells you to choose the best prospect given the probabilities and the values of each option, and a principle that allows you to choose any option that you know would be permissible if you had all the relevant information. Although Hare favors prospectism, he concedes that he has no conclusive argument against deferentialism. I think Hare's characterization of the choice situation in terms of the decision theory framework obscures some features of decisions with options on a par. To my mind, there is no question that Box B is the only rational option here. Imagine trying to justify a decision to take Box A. Suppose someone said, "I took Box A because it might have had something that I would have chosen over the contents of Box B even if I were to get an extra \$100." However the same is true of Box B, except that if you choose Box B, you get an extra \$100. Nonetheless I agree that we need an explanation of why my knowledge that my better-informed self might choose Box A is not a consideration that permits me to choose Box A.

C10P40 Let us look at my choice when I know that Box A has the Fabergé egg, and Box B the wedding album, and neither option affords me an extra \$100. I appreciate the unique beauty and the historical significance of a Fabergé egg. This is clearly sufficient reason to choose the egg. On the other hand, the wedding album's value has a radically different source; its relation to my life

with my beloved, as well as the memories of that wonderful day, also constitute sufficient reason to choose the wedding album. In such a case if I choose Box A, I choose it *because it contains the Fabergé egg* (or perhaps, because it instantiates the value-constituting properties of a Fabergé egg). I'm not indifferent between these two options; we should not represent this choice as a comparative choice between two values whose quantitative nature is such that it permits me to choose one or the other. Rather, the *fact that Box A contains the Fabergé egg is a good reason to choose Box A* and this suffices to make choosing Box A permissible, as long as there is nothing about Box B that defeats this reason. Of course, the same story *mutatis mutandis* would explain the permissibility of my choosing Box B if I were to choose Box B. And, on the assumption that the relevant facts are known, adding \$100 to either option is not enough to change any of these facts.

C10P41 It is worth noting that in the case in which I know what each box contains and I can't make up my mind, it is permissible to randomize my choice by, say, flipping a coin. Moreover, in the absence of a coin to flip, I could use various procedures in lieu of a more strictly randomizing procedure; there would be nothing wrong in picking the property of "being the longer box" as settling which one I choose. But it would be wrong to think that my reason to choose Box A is that it is the one that contains the Fabergé egg and it is the longer box; in fact, if this were the reason that would be something wrong with how I was making my choice. And this can be seen by the fact that "being the longer box" cannot play the same role in the case in which we sweeten the opaque box. If it is rationally impermissible to pick Box A in such a case, it is also impermissible to pick it once I realize that Box A is the longer box. Relatedly, what allows me to rely on the *Pareto Principle* in the case of ignorance is not that I have no way to pick out uniquely the object that is in Box A (which I know to be on a par with object B). In fact, the description "The object in Box A" picks out the Fabergé egg and is available to me. But since 'being in Box A' is irrelevant to my reasons for choosing either option, my choice situation is still the one described above.

C10P42 In sum: In the scenario in which I know what each box contains, the reasons to choose Box B do not defeat the reasons to choose Box A (or vice-versa). But the same is not true if I do not know what is in the boxes. I cannot choose Box A because the Fabergé is there, because that fact is not accessible to me. I could choose Box A because there *might* be a Fabergé egg there, but this possibility is there on Box B too. With respect to the Fabergé egg and the wedding album, each box offers me the same reason to choose it: it

might contain (it has .5 chance of containing) a Fabergé egg; it might contain (it has .5 chance of containing) the wedding album. Box B on the other hand offers something that Box A does not offer: \$100. So Box B is favored by an *Extended Pareto Principle*:

C10P43 In situations involving a choice between φ -ing and ψ -ing, where φ -ing does better than ψ -ing in realizing at least one value, and does better or equal in realizing all other values involved in the choice, a rational agent chooses to φ , if she chooses to φ or ψ .

C10P44 It should be clear by now why the fact that my better-informed self would be permitted to choose Box A carries no weight: I simply have no access to the reason they have when making their decision. The permissibility of an action, at least its subjective permissibility, depends on which reasons are accessible to me, or, in other words, which reasons I can act on.¹³

C10S6

6. Kant to the Rescue

C10P45 A similar reasoning applies to **Basic Case**, or so I will argue now. Suppose I am told that I have the option to save one stranger's life (A) or the lives of two strangers (B and C). Each person in the universe has dignity and if I was saving one of them, the *No Defeat Principle* would apply and I would not be required to abandon saving the one for the sake of saving the two. But this is not my situation when I am deciding what to do. I know nothing of the persons I might save and thus my position is a choice in which each person that I would save could be anyone, but in one option I get to save one person who could be anyone (just like in the choice of one person), and *also* an opportunity to save another person who could also be anyone, except not a person identical to the first one. In other words, the choice is between saving *a* person and saving a person *plus yet one more person* (See Table 2 for a rough version of the choice). I do have available to myself descriptions that

¹³ I am assuming that there is a legitimate notion of subjective permissibility (whether or not there is also a notion of objective permissibility), both in the case of morality and in the case of instrumental rationality, and that this is the relevant notion from the point of view of the agent making a decision. Even if one were to reject such a notion, one would still need to say something about principles that guide rational or virtuous agents making decisions under limited information, and these are the principles that are relevant here.

C10T2 **Table 2:** Choice Among Strangers

Option A	Option B
$1/n(\text{person}_1, \dots, \text{person}_n)$	$1/n(\text{person}_1, \dots, \text{person}_n); 1/n - 1(\text{person}_1, \dots, \text{person}_{n-1})$

will successfully pick out person A (like “the person who I can save but only singly”), but since these predicates are irrelevant for my reason to engage in either action, they do not change my choice situation.

C10P46 But exactly what is relevant? Well, dignity is a value a person has, on Kant’s view, simply in virtue of being a rational agent. So it seems that no property of the person will be relevant for my reasons to save them, and we would end up with the surprising conclusion, if welcomed by some, that the *Pareto Principle* always applies and one should always save the many over the few. However, the matter is not so obvious. First, there are at least some people who stand in certain relations to me that are obviously relevant and that can potentially provide reasons that defeat my reason to save *a* person. As we said above, partiality (or liability) might provide reasons to save my relation (or the non-liable person) that defeat the reason to save *a* person.

C10P47 More importantly for our purposes, dignity attaches not to a generic property or to a trope, but to the *particular person* who has the property in question. Though a particular agent has intrinsic value in virtue of possessing this property, the value attaches to them in particular, rather than to the state-of-affairs such that the property of personhood is instantiated. So as long as my reason to act picks out *the particular person* I am about to save rather than *a* person, the *Pareto Principle* will not apply to it, and my reason to save *this person* will be undefeated by my reason to save the many. In the case of baby Chris, this is exactly what happens: as Sam is already engaged in saving baby Chris, her thought is already directed at *baby Chris himself* and thus her saving him is undefeated by the opportunity to save two babies. This would be true even before she enters the water: at the moment she sees baby Chris she has settled on saving baby Chris himself, and her reason is a reason to save this person in particular. In fact, had she received the update just as she started going in the direction of the pond, but before she makes any acquaintance with baby Chris, it would be difficult to justify not turning around and saving the two children instead. And even if you thought that a mere acquaintance is not enough to generate a proper

reason of partiality to save David, the barista of my local cafe,¹⁴ in another scenario this might end up sufficing to allow me to *continue* in my endeavor to save David. For instance, suppose I were at a beach and learned that a large ship had sunk and people needed to be saved. This is a dangerous task, and there are already enough people who are taking it upon themselves to save those shipwrecked, so I decide not to take the risk. But then I spot David barely keeping himself afloat, and I think “I need to get David to the shore”. As I swim towards David, I see that I can save two people instead if I turn in the other direction. Am I allowed to continue and save David? I think so.

C10P48

Of course this reasoning can be extended significantly beyond cases in which I am already in the midst of, or even settled upon, a particular course of action. After all, if what makes a difference is that the person herself appears in my reason, the person could also appear in my reason at the time of the decision. If I decide to save *David himself*, my reason is undefeated by the prospect of saving two people. This might appear counterintuitive, but I think it is a welcome result. Let us go back to our scenario of David being one of the victims of a shipwreck. Let us assume that at first I do not join the rescue efforts, and let us assume that even though this will result in fewer people being saved (there are not nearly enough people on the beach to save everyone who is drowning), it is permissible for me not to join (even though I am certain to save those whom I attempt to rescue, there’s a significant chance that I’ll die as a result of it due to my fish allergies). But then I remember that David was in the ship. I also know he was the only single person in a cruise for newlyweds (he was the hired barista for the cruise) and that he’ll be the only person by himself on the sea (the newlyweds keep very close to each other and far from others, and this configuration remained after the shipwreck). Since each person on the beach will have time to go only on one rescue mission, they are all facing a situation like **Basic Case** (except that there is more than one set of two people that each of us can save instead of saving David), and thus it is quite clear to me that David will not be saved. So, I remember the times when David and I were bantering while he was serving me coffee and I think to myself “I’ve got to save David”, and I then jump into the water in order to save him. It doesn’t seem that I act badly now when I choose to save David even if I could save two other people instead. Now, this is a case of supererogation. But, first, in other cases of supererogation, I seem to be under the same obligations as in an “obligatory”

¹⁴ Following Setiya’s suggestion mentioned above in footnote 8.

rescue case *once I decide that I'm making the necessary sacrifice*; if as I am saving David I see that I can easily save another person with no extra effort, I would be then obligated to save the other person as well. Moreover, it is not hard to think of similar cases in which I have a clear duty of rescue; in fact, I do not think it would make any difference to our verdict if I had no allergies (and was thus obligated to rescue someone), and, I was about to join the rescue efforts, I realized that, given the configuration of the newlyweds, unless I were to save David, David would not be saved..

C10P49 Note that I am not claiming that “being a particular person” is a morally relevant consideration; after all, what kind of consideration this would be?¹⁵ Rather, what is relevant for this case is the fact I am saving *David in particular* is a different kind of reason than the fact that I am saving *someone*. Taurek thinks that if weak forms of partiality can ground duties of partiality that override the duty to save the many, the duty to save the many cannot be very stringent. But on our account, the issue is not that some kind of partiality, a mere casual or even newly acquired acquaintance, overrides the value of saving an extra life; rather, having access to the reason “saving *David himself*” makes the *Pareto Principle* inapplicable.

C10P50 But what exactly makes it possible for me to have a reason to save *this particular person*, rather than just a reason to save *a person*, or, to put it slightly more controversially, when is it possible for *person herself* to figure in the content of my practical reasons? At the very least the agent must be able to pick out the person, so the agent must be at least capable of having singular thoughts about them. But this might not be sufficient to allow a particular person to figure in my reason for action. Suppose I catch a glimpse of Jeri at a distance for a few seconds. Can the thought “*Jeri is drowning*” figure as a reason to save Jeri that is relevantly different from the reason given by the thought “someone is drowning”? It seems that the answer is “no”, and perhaps practical acquaintance requires more; my own view is that practical acquaintance requires a more robust engagement with my practical thought. Cases like Chris’ and David’s involved much more substantive agential engagement with the rescue target than a mere passing glimpse. Of course, this is just a promissory note, and to make good on it we need a more precise account of “robust engagement with my practical thought”. But however we account for what makes this kind of practical acquaintance

¹⁵ But see Setiya (2023). Setiya talks about “personal acquaintance”, which overlaps, but does not perfectly match, what I call “practical acquaintance” below. Setiya does think that personal acquaintance is morally relevant in similar cases.

232 SERGIO TENENBAUM

possible, I have argued here that it makes available to us a reason that changes the character of our choice situation.

C10P51 This observation lets us also answer a potential concern: namely, that the structure of dignity makes **Basic Case** radically different from instances of “sweetening in the dark”. We already know that each person has dignity; while each box in Hare’s example is literally just a vessel for the different values it carries, human beings are not just vessels of some unknown value. Couldn’t I just permissibly take stream A so as to save someone who has dignity, no matter who that person turns out to be? Indeed, this is a good reason to take stream B, but this is also the very same reason I have to save the “first” person in stream A. It is exactly the fact that persons are not mere vessels of dignity that makes **Basic Case** structurally similar, in the relevant way, to the cases of sweetening in the dark. Note that in Hare’s example, I do know that Box A contains something whose value would give me a sufficient reason to choose it (or, to put it somewhat controversially, I know that the value of Box A’s contents is incommensurable with the value of Box B’s contents). However, this reason is not accessible to me; I can only act on the reason that the box has either the Fabergé egg or the wedding album. Similarly, each person has dignity, and thus I know that if I steer to stream B there will be someone there whose value would give me sufficient reason to save them. But we are not just vessels of dignity; thus the reason to save this *particular* person is not available to me. I can only act on the reason that I can save *a* person here. This reason, however, is *identical* to the reason I have to save Larry (or Mary) and thus the *Pareto Principle* kicks in and forces me to go to stream A.

C10P52 It is also worth noting that *how* I involve others in my agency can be morally relevant whether or not I am practically acquainted with them. Doubtless, if I find a way to push a person off a footbridge in order to save five without gaining acquaintance with my victim, I am still using this person as a means. But nothing in my argument jeopardizes this platitude. You cannot treat anyone merely as a means, and so you cannot treat someone as a means even when you don’t know who that person is. My reason for saving the person is defeated by the prohibition to use *any* human being in this manner. Here the comparison with cases of incommensurability might also be helpful. If someone offers to pay me \$60 for an object, and all that I know about this object is that its value is incommensurable (or on a par) with a Fabergé egg, I can easily conclude that I must turn down the offer. After all, my reason to make \$60 is defeated no matter what the object turns out to be. What follows from my argument,

both in the “prudential” and the “moral” case, is that ignorance (or lack of practical acquaintance) potentially changes the nature of my reasoning, not that it changes the nature of the value in question.

C10P53 Finally, it is worth noting that this kind of reasoning, based on the Pareto Principles above, might have further implications. For instance, it can help the Kantian, and other deontologists, explain the plausibility of a different kind of verdict, in a different context. Let us look briefly at a consideration favoring effective altruism: whether or not you are required to donate to global charities, if you do decide to donate to these organizations, you should donate to the more effective ones. This kind of consideration seems overwhelmingly plausible if we are deciding between two charities investing in similar efforts, but one of them is significantly more effective than the other. If “Against Malaria” and “Somewhat Against Malaria” are both fighting malaria but the former saves many more people than the latter with the same amount of money, we seem to have decisive reason to prefer the former even if the people saved are not the same. However many would want to resist this conclusion if the charities are committed to different issues (say, one of them is fighting malaria while the other is trying to help refugees), or even if I am particularly committed to those who live in the “Somewhat Against Malaria” area (an area not covered by “Against Malaria”). In such cases, it seems plausible that I do no wrong if I prefer the less efficient charity. Non-consequentialists might want to explain the permission to donate to “less efficient” charities in terms of the fact that the claims of different people cannot be compared in the way suggested by the effective altruist, but this leaves us wondering why the same reasoning does not apply to the first type of case. Perhaps the Extended Pareto Principle (or some similar principle) grounds the different deontic verdicts in these two types of cases.

C10P54 So can Kant count? Well, no, but Kant can do all that counters can and much that counters can't.¹⁶

C10S7

References

- Chang, R. (2002). The possibility of parity. *Ethics*, 112(4):659–88.
Hare, C. (2010). Take the sugar. *Analysis*, 70(2):237–47.

¹⁶ For very helpful comments, I owe many thanks to Sarah Buss, Kyla Ebels-Duggan, Thomas Hurka, Jennifer Nagel, Julia Nefsky, two anonymous referees, and audiences at the 2021 Workshop in Normative Ethics and the Department of Philosophy at Bilkent University.

- Horton, J. (2018). Always aggregate. *Philosophy & Public Affairs*, 46(2):160–74.
- Igneski, V. (2006). Perfect and imperfect duties to aid. *Social Theory and Practice*, 32(3):439–66.
- Kant, I. (2012). *Groundwork of the Metaphysics of Morals*, trans. Mary Gregor. Cambridge University Press.
- Parfit, D. (1978). Innumerate ethics. *Philosophy and Public Affairs*, 7(4):285–301.
- Saunders, B. (2009). A defence of weighted lotteries in life saving cases. *Ethical Theory and Moral Practice*, 12(3):279–90.
- Setiya, K. (2014). Love and the value of a life. *Philosophical Review*, 123(3):251–80.
- Setiya, K. (2023). Other people. In Buss, S. and Theunissen, N., editors, *Rethinking the Value of Humanity*. Oxford University Press.
- Singer, P. (1972). Famine, affluence, and morality. *Philosophy and Public Affairs*, 1(3):229–43.
- Steuwer, B. (2021). Aggregation, balancing, and respect for the claims of individuals. *Utilitas*, 33(1):17–34.
- Taurek, J. M. (1977). Should the numbers count? *Philosophy and Public Affairs*, 6(4):293–316.
- Tenenbaum, S. (2017). Action, deontology, and risk: against the multiplicative model. *Ethics*, 127(3):674–707.
- Thomson, J. A. (2022). Relief from rescue. *Philosophical Studies*, (179):1221–39.
- Timmerman, T. (2015). Sometimes there is nothing wrong with letting a child drown. *Analysis*, 75(2):204–12.
- Tomlin, P. (2017). On limited aggregation. *Philosophy and Public Affairs*, 45(3):232–60.