One to One

[8,770 words]

Common-sense morality recognizes that the numbers can matter. If we can build a public transit system in one or another city, the fact that more people will use the system in one city surely bears on where to build it. The number of beneficiaries can affect what we ought to do.

But common-sense morality also recognizes cases where the numbers do not matter. If a stagehand faints just before the curtain rises, it would be wrong to allow the play to continue just to avoid inconveniencing each of many viewers with a slight delay. We must help the stagehand immediately, no matter how many are waiting for the show to begin.[1] Here, the number of beneficiaries does not affect what we ought to do.

Many theories explain these verdicts by ascribing to morality both an aggregative component, which explains why the numbers matter in the transit case, and an non-aggregative component, which explains why the numbers do not matter in the theatre case.[2] But positing two components to explain two kinds of cases raises concerns about overfitting and theoretical unity, while appeals to aggregation raise familiar controversies. This paper proposes a theory that uses one mechanism and no aggregation.[3]

The theory takes inspiration from a basic form of moral reasoning where we (i) take up another's perspective, (ii) recognize how that person's interests compare to our own, and (iii) modify our own attitudes accordingly. I will say more about what exactly this involves, but I hope the general idea is familiar from experience. A couple bickering about how to arrange

---

[1] Scanlon 1998: 235.

[2] The leading such theory is Voorhoeve 2014's Aggregate Relevant Claims, which is further developed by van Gils & Tomlin 2020, Hart 2022/2024, Mann 2021, Rüger 2020, Steuwer 2021, Tadros 2019, Voorhoeve 2017, and others.

[3] Many papers have similar aims; footnote 14 compares the theory of this paper to others.

the new apartment may need to pause, take up the other's point of view, and accept some compromise. Let us stipulatively call this whole process the "internalization" of another's perspective.

The gambit of this paper is to use only facts about how individuals internalize the perspectives of other individuals to explain cases like those above. To that end, the paper describes internalization in greater detail (§1) and defends four principles that govern it (§2). These principles imply a theory that considers how individuals might internalize the perspectives of others and then requires one to minimize the maximum degree of unacceptability to any individual. On this theory, the numbers matter when deciding where to build a transit system but not when deciding whether to save the stagehand (§3). The paper concludes by addressing an objection (§4).

§1: What Is Internalization?

This section answers the titular question, hopefully revealing "internalization" to be just an ugly, unfamiliar name for a familiar and compelling pattern of moral reasoning.

Start with an illustration. Suppose you accidentally bought an extra movie ticket. Your friends, Alice and Bob, each want the ticket. It would be unreasonable, though, for either of them to insist on receiving the ticket. If one of them insisted that you must give them the ticket, you could answer by asking them to consider the other friend's perspective. They should then recognize that both of them have similar claims to the ticket, that the other friend matters just as they do, and thus that they cannot reasonably insist on receiving the ticket. They must, accordingly, drop their insistence that you give them the ticket.

We can distinguish three stages in this example:

1. The pushy friend takes up the perspective of another
2. The pushy friend compares the other's perspective to their own
3. The pushy friend modifies their attitudes in response to this comparison

We discuss each in turn.

First, taking up the perspective of another involves seeing things from their point of view, where this includes their outlook, values, tastes, history, dispositions, and so on. A litigation lawyer raised in 1960s Alabama will see things differently from a programmer working in 2020s Hangzhou.[4] In principle, people's perspectives are incredibly complex objects, but for practical purposes one can often ignore many of the details. And people's perspectives may conflict with applicable moral and non-moral standards; for example,

[4] For helpful discussion, see Goldie 2002: 194 – 205.

someone may believe that animal pain has no moral significance, or that animals do not feel pain. Taking up their perspective nonetheless involves taking on this outlook, warts and all.

Second, comparing another's perspective to one's own involves considering what is at stake from their perspective, what is at stake from one's own perspective, and comparing these by applicable moral standards. Each perspective assesses different actions in different ways, but not all of these assessments pass muster. Someone may assess the prohibition on loud noises after midnight to be a violation of personal liberty that far outweighs their neighbors' interests in a good night's sleep. Their neighbors may have the opposite inclination. Comparing one's perspective to another's perspective does not involve comparing them by one's own lights or by the other's lights, but rather by the moral standards that actually govern.[5]

Third, modifying one's attitudes in response to this comparison involves changing one's assessment of possible actions to reflect the importance of those actions from the perspective of another. The pushy friend, for example, after recognizing that the other friend's interests are as important as their own should drop their insistence on receiving the movie ticket. They should accept that you may randomly pick one of them to receive the ticket, give it to the friend you see less often, or decide with some other criterion. That is the appropriate response to recognizing that one's interests matter just as much as those of someone else.[6]

Sometimes, one person has more at stake than another. The exhausted neighbors have more at stake than the opera buff who wants Rigoletto's anguish to resound through the building (and who keeps looping *Se t'involi, qui sol rimarrei*). The opera buff, after comparing

---

[5] For helpful discussion, see Kamm 1993: 153 – 154.

[6] Scanlon 1998: 196 – 197.

their perspective to that of their neighbors, should agree to dial it down and should find their prior conduct unacceptable. But even here, the perspective of the obnoxious opera buff should still mean something to the neighbors. A frustrated neighbor who compares their plight to that of the opera buff should still temper their reaction somewhat; the opera buff has something at stake and recognition of this fact should make their actions ever so slightly more acceptable to the neighbor. Even after modifying their assessment to reflect the importance of the music to the opera buff, though, the neighbor may continue to find their conduct highly unacceptable and to insist that the opera buff turn it down. Taking another's perspective into account requires that one's assessment of an action reflect what is at stake for them, not that they must get their way. The perspective of another can thus affect, and sometimes to only a small degree, how unacceptable one finds some action, without altering whether one finds that action unacceptable.

Let internalization name this process of taking up another's perspective, comparing that perspective to one's own, and modifying one's assessment of various actions accordingly. Internalization is a success term; mere attempts to take another's perspective into account do not qualify. And the name may be jarring, but I hope that the process it denotes resonates with our own experiences of moral deliberation.

Two remarks about the role internalization plays in the theory to come. First, internalization does metaphysical, not just epistemic, work. In other words, facts about internalization partly determine what one should do and do not merely provide evidence about this. The fact that the pushy friend, after internalizing the other friend's perspective, would not insist on receiving the ticket *makes it the case that* you need not give them the ticket. The former fact does not merely indicate the latter.

Second, the theory's prescriptions depend on what would be true *if* people internalized others' perspectives, not on whether people actually do so. You are permitted to give the movie ticket to the laidback friend regardless of whether the pushy friend actually internalizes their perspective. (Perhaps the pushy friend simply refuses to believe that the other friend, or anyone else, could possibly care as much about the cinema as they do). The grounds of this permission are the attitudes the pushy friend would have were they to internalize the other friend's perspective, not their actual attitudes.

Relatedly, the theory does not require people to actually internalize the perspectives of others. According to the theory, what an agent ought to do depends on the attitudes people would have if they internalized the perspectives of others, but this does not imply that those people actually ought to internalize those perspectives. To draw an analogy, whether I ought to attend an estranged relative's funeral may depend on whether this would make others uncomfortable; this does not imply that they ought to feel uncomfortable. Internalization is a mental action and as such, there are often reasons for and against its performance.[7] The theory places no thumb on the scale here.

[7] See Bailey 2021, 2022.

§2: Principles for Internalization

There is a family of moral theories that use only internalization. For example, one might think that an action is permissible iff the tallest person, after internalizing the perspective of the shortest person, finds it acceptable. Less fancifully, one might think that an action is impermissible iff no one could find that action acceptable after internalizing the perspective of the person who stands to lose the most from it. On this theory, those with the most at stake are morally decisive.[8]

This section defends four principles about internalization that jointly characterize a particular theory. I accept these principles and therefore take this to be an argument for the theory, but those who doubt these principles can still use them to organize conceptual space, develop new theories, and pinpoint my errors.

The first two principles constrain which ways of asking individuals to internalize the perspectives of others can bear on what we ought to do. Some are not possible, e.g. asking someone to internalize the perspectives of multiple others, while others are unfair, e.g. asking everyone to internalize the perspective of the same person. These principles exclude such possibilities, constraining which patterns of internalization are eligible to bear on what we ought to do. The first principle requires that no one internalizes the perspective of multiple individuals (No Multitudes) while the second principle requires that no one perspective is internalized by multiple individuals (No Multiplicity). The third principle accepts as eligible all remaining patterns of internalization (Sufficiency).

The fourth principle connects the eligible patterns of internalization with what we ought to do. It states that we ought to minimize the maximum degree of unacceptability to any

[8] This resembles minimax complaint, which requires one to minimize the strongest individual complaint. For helpful discussion, see Reibetanz 1998: 300 and Horton 2017: 55.

individual under some eligible pattern of internalization (Minimax). What exactly this means will become clearer as things proceed.

§2.1: No Multitudes

Internalization relates one individual to another; it asks someone to take up another individual's perspective, to compare perspectives, and to modify one's attitudes in response. As such, one cannot internalize the perspectives of multiple others. The opera buff can take up the perspective of one neighbor and compare that perspective to their own. But the opera buff cannot take up the perspectives of all the neighbors or compare all of these perspectives to their own.

To see this, consider what it would mean for the opera buff to take up the perspectives of all the neighbors. The building may house new parents who fear the noise will wake their finally sleeping children, and it may house teenagers who want to play their video games in peace. Does taking up both perspectives at once involve imagining a teenage new parent who wants their children to stay asleep so they can play video games late into the night? Or does it involve imagining someone uncertain about whether they are a new parent or teenage gamer? Neither of these corresponds to the perspective of any actual neighbor, so neither offers a way to take up every neighbor's perspective.[9] There is no perspective we can take in order to take up the perspectives of every neighbor. As a result, there is no such perspective for the opera buff to compare to their own. Internalization breaks down.

---

[9] Even if every neighbor had qualitatively identical perspectives, taking up one neighbor's perspective would not be the same as taking up all of their perspectives simultaneously. A child who complains that their parent rarely see things from their point of view would not be assauged by learning that their parent often sees things from their twin's point of view, even if the two points of view are qualitatively identical. Taking up someone's perspective differs from taking up a perspective qualitatively identical to theirs.

Individuals cannot internalize the perspectives of multiple others, so no eligible pattern of internalization can ask any individual to do so. We can state this as a principle:

No Multitudes

In any eligible pattern of internalization, no individual internalizes the perspectives of multiple others.

where the name reflects how internalization relates individuals to individuals, not multitudes.

For some, this principle will raise doubts about the importance of internalization for moral theory. One may think, "If we can compare what one person has at stake to what another person has at stake, then surely we can compare what two people together have at stake to what another person has at stake. If 'internalization' does not include the latter process, then internalization fails to carve morality at its joints and should not be central to our moral theories."

I deny the conditional: we can compare what one person has at stake to what another person has at stake, but we cannot similarly compare what two people together have at stake to what another person has at stake. This may seem hard to believe, and a full defense would require a separate paper, but consider the other dimensions along which we sometimes compare individual perspectives.[10] We might delight in someone's curiosity, marvel at their insightfulness, or match their cynicism. Recognizing that someone is more curious than I am involves comparing my perspective to theirs along the dimension of curiosity and finding that theirs exceeds mine. But this does not imply that I can also compare myself to a group of people along this dimension. I cannot, for example, compare my curiosity to that of a class of undergraduates, even if I could compare it to that of each undergraduate. Many properties that

[10] Taurek 2021: 312 – 314.

only perspectives can have allow for comparisons between perspectives but not between groups of perspectives. If only perspectives can have the property of mattering morally, we may expect this conditional to fail.

§2.2: No Multiplicity

Internalization involves modifying one's attitudes to reflect the importance of someone else's perspective. It involves taking someone else's perspective so seriously, one changes one's own assessment of the actions. Given this, certain patterns of internalization can seem unfair.

For example, it would be unfair to ask everyone to internalize my perspective. Why me? As a result, the fact that an action would be acceptable to everyone if they each internalized my perspective does not show that the action is permissible. The pattern where everyone considers my perspective is not eligible to bear on what we ought to do.

The obvious problem with the above is that it arbitrarily gives someone, namely me, a special moral status. But even without the arbitrariness, I argue, there is something inherently wrong with the status: no perspective deserves to be internalized by multiple others. It is unfair to ask multiple people to internalize the same person's perspective, and as a result, such patterns of internalization do not bear on what we ought to do. We can state this as a principle:

No Multiplicity[11]

[11] Instead of requiring that no one's perspective is internalized by multiple others, one might require that no *part* of anyone's perspective is internalized by multiple others. The idea would be to allow multiple people to internalize the same person's perspective, provided they each internalize separate parts of that perspective. For example, perhaps the construction outside my apartment both wakes me up and distracts me during the day. One person might be asked to consider how it wakes me up while another person might be asked to consider how it distracts me during the day.

I like this thought, and it largely comports with what I say in the main text. But as it raises further questions about the metaphysics and mereology of perspectives, I do not discuss it further here.

> In any eligible pattern of internalization, no individual's perspective is internalized by multiple others.

Why is it unfair to ask multiple people to internalize the same person's perspective? Fundamentally, this is because any action, in the sum total of its many and variegated repercussions, affects us once. However the action harms us, it harms us once; however it blesses us, it blesses us once. After someone internalizes our perspective, coming to grips with all we stand to gain or lose and even placing this on level footing in their outlook with what they themselves have at stake, we have been fully taken into account. We cannot ask for more. We could not, for example, insist that a second person and a third and so on similarly take up our perspective and weigh it next to theirs in their outlook. Each counts for one. That is why one cannot count for many.[12]

This is the metaphysical explanation, providing reasons that *make* it unfair to ask multiple people to internalize the same person's perspective. There is also an epistemic reason to believe this: allowing multiple people to internalize the same person's perspective leads to the absurd consequence that we should flip a coin when deciding whether to give some benefit to a greater or lesser number of people.

For example, suppose that I can screen a rare film on either Monday or Tuesday. Of my dozen friends, Alice is only free on Monday, while the others (Bob, Carol, ..., Lucille) are only free on Tuesday. If there is nothing wrong with asking multiple people to internalize the same person's perspective, then I could apparently reason as follows. Bob initially favors Tuesday,

---

[12] The idea that we each deserve to be taken into account once because we each have one life to lead differs from Kamm's Balancing Argument, which holds that equality requires an additional person's similarly weighty interests to make a difference by breaking ties among otherwise permissible options (Kamm 1993: 116 – 119, see also Scanlon 1998: 232). For one, the former applies even when people have disparate interests at stake, since it remains true that each will suffer such gains or losses from the action only once. For another, only the latter relies on the thought that people should make a difference to what is permissible. As others note, this idea does not obviously support the intended conclusion (Raz 2003: 362).

but I could ask Bob to consider Alice's perspective, realize they have similar interests at stake, and thereby come to insist that I flip a coin to decide. Carol also initially favors Tuesday, but I could ask Carol to consider Alice's perspective, realize they have similar interests at stake, and thereby come to insist that I flip a coin to decide. And so on through Dennis, Edie, Fannie, George, Henry, Isidore, Jocie, Kevin, and Lucille. In short, I could ask everyone to internalize Alice's perspective (while asking Alice to internalize, say, Bob's perspective), and then everyone would insist that I flip a coin. If this pattern of internalization bears on what I ought to do, then I am permitted to flip a coin.

But both the conclusion and this reasoning are absurd. Some evidence: I would never deliberate in this pattern, and I would be embarrassed if my friends realized what I was doing. Moreover, I imagine a disinterested observer would criticize me as follows: "You are asking each of Bob, Carol, ..., Lucille to take Alice's perspective as seriously as they take their own. But there are eleven of them and one of Alice. It is as though you think Alice has eleven lives that each deserve to be weighed against each of theirs. But she doesn't. She's one person. Asking all of them to take up Alice's perspective treats what Alice has at stake out of all proportion to what she actually has at stake." These thoughts support No Multiplicity.

§2.3: Minimax

No Multitudes and No Multiplicity restrict the eligible patterns of internalization. I assume that they are the only such restrictions, so that any pattern of internalization that satisfies their strictures may bear on what we ought to do.[13] In other words, when some way of asking

[13] Additional restrictions may be plausible (e.g. as many as possible internalize the perspective of another), but I am aware of none that would change the results to come. It may be that such restrictions are not fundamental but rather derived from these principles.

individuals to internalize the perspectives of others satisfies the strictures of No Multitudes and No Multiplicity, then this is one way we might justify choosing some action. Nothing rules it out in advance. Stated as a principle:

Sufficiency

Any pattern where (i) no one internalizes multiple people's perspectives, and (ii) no one's perspective is internalized by multiple people's perspectives is eligible.

Typical choices allow for many eligible patterns of internalization. For example, the obnoxious opera buff might internalize the perspective of one of the new parents. Or their child. Or one of the teenage gamers. Any of these patterns would satisfy No Multitudes (as no one internalizes multiple perspectives) and No Multiplicity (as no one's perspective is internalized by multiple people). These patterns can differ, though, in how people ultimately view the various options. Suppose the noise matters more to the new parents than to the teenage gamer. On one eligible pattern, the opera buff internalizes a parent's perspective and finds blasting *Rigoletto* highly unacceptable; on another, the opera buff internalizes a teenage gamer's perspective and finds it only somewhat unacceptable.

Minimax connects the eligible patterns of internalization with what we ought to do. The idea is to consider, for each action, the best case one could make for its performance. We consider the best case for performing each action because, in general, the fact that there are bad arguments for choosing something does not count against its selection; what matters is the strength of the best arguments. On any eligible pattern of internalization, different people will find the action acceptable or unacceptable to different degrees. As we have seen, how an individual assesses an action can depend on whose perspective they internalize. On different

patterns of internalization, different people will have different attitudes towards the action under consideration.

Imagine rifling through all eligible patterns of internalization to find one that minimizes the maximum degree of unacceptability to anyone of that action. This pattern of internalization makes the best case possible for that action. We then rank actions according to their maximum degree of unacceptability to anyone given the best case for that action. We may pick any action that minimizes this degree. In brief, we may choose any action that, under its best case pattern of internalization, performs no worse than any other action under its best case pattern. Stated as a principle:

> Minimax
>
> One may choose all and only the actions that minimize the maximum degree of unacceptability to any individual when compared to each alternative on the pattern of internalization most favorable to the former action in this comparison.

The aim of minimizing the maximum degree of unacceptability to any individual emerges as the natural choice given our committment to internalization. Internalization allows an individual to compare what they have at stake to what another individual has at stake, but it does not allow one to sum the degrees by which various people individually find an action unacceptable. This reflects an outlook on which individual perspectives matter morally, but groups or sums of individual perspectives do not similarly matter.

Given this, the idea of minimizing the aggregate amount of unacceptability to any individual seems out of place. And while we could, say, maximize the maximum degree of acceptability to any individual, or maximize the number of people to whom the action is acceptable, none of these seem like comparably natural objects of our moral concern. The

person to whom an action is least acceptable occupies a special place in our moral thought, as Minimax acknowledges.

§3: From Principles to Theory

Our theory is the conjunction of No Multitudes, No Multiplicity, Sufficiency, and Minimax.[14] To understand this theory, it will help to see how it resolves some cases, including the transit and theatre cases with which we began.

§3.1: Transit

For ease of illustration, suppose that building the transit system in Einhaus will benefit one person, Una, while building the transit system in Twyford will benefit two people, Dora and Tressa. Each stands to benefit to a similar degree.

---

[14] How does the theory relate to other views?

One family of views uses both aggregative and non-aggregative components. Within this family, some views (e.g. Voorhoeve 2014) aggregate deontic phenomena, such as the claims of individuals, while others (e.g. Scanlon 2021, Zhang 2024) restrict aggregation to axiological phenoemena, such as the value of wellbeing, but allow such components to bear on non-aggregative components, such as whether an individual can reasonably reject some principle or whether some action treats an individual with equal consideration. The theory of this paper differs from these by rejecting even axiological aggregation. (A fortiori, the theory departs from the family of views that use only aggregative components.)

Another family of views uses only non-aggregative components. Philosophers disagree about exactly which views manage to do so, but the following surveys some plausible candidates. First, some views (e.g. Sobjectivity 1 from Kamm 1993: 154 – 163, Taurek 1977) imply that the numbers do not matter, but the numbers can matter on the theory of this paper as §3.1 shows. Second, some views (e.g. Gustafsson 2021, Hirose 2001) use transitivity along with apparently non-aggregative principles to derive certain conclusions, but the theory has cycles so cannot accept transitivity. Third, some views (e.g. Henning 2023: 771, Scanlon 1998: 239 – 240) use the idea of some claims being 'relevant' or 'irrelevant' to other claims; the theory does not use this distinction. Fourth, some views (e.g. the Context-Aware View from Kamm 2006: 58, Kumar 2011: 143) use the idea of someone's claims 'balancing' or 'neutralizing' another's claims, but they restrict this to cases where people have similar interests at stake. The theory uses the structurally similar idea of internalization but applies it even in cases where people have very different interests at stake. This yields some distinctive implications, as §3.4 discusses.

By No Multitudes, No Multiplicity, and Sufficiency, there are thirty-four eligible patterns of internalization.[15] But we can ignore most of these because, by Maximin, what we ought to do depends only on the best case that can be made for each action.

Imagine building the transit system in Einhaus. This frustrates Dora and Tressa while benefiting Una. The best case for this choice involves asking Dora (or Tressa) to internalize Una's perspective. This may appease Dora, but it still leaves the action unacceptable to Tressa.

Alternatively, imagine building the transit system in Twyford. This frustrates Una but benefits Dora and Tressa. The best case for this action involves asking Una to internalize Dora's (or Tressa's) perspective. This appeases Una, rendering the action acceptable to all.

Thus, building the transit system in Twyford minimizes the maximum degree of unacceptability to anyone. By Maximin, we must build in Twyford. Similar reasoning applies to any situation where one can bestow sufficiently similar benefits on a smaller or larger group.

§3.2: Theatre

Now consider the theatre case. Suppose Una is the stagehand who has just fainted and there are a million audience members waiting impatiently. Again, there are many possible patterns of internalization, but we can focus on the best case for each action.

Imagine helping Una immediately. This inconveniences the audience members but ensures that Una receives prompt medical attention. We can ask one of the audience members to consider Una's perspective, leading them to accept and even insist on helping her. However, this action remains mildly unacceptable to the other audience members.

---

[15] See Fact 1 of the appendix.

Alternatively, imagine waiting to help Una. This puts her health at risk while helping the audience members. We can ask Una to internalize the perspective of an audience member, but even after doing so, she can insist on receiving help immediately given how much more she has at stake than the audience member. Allowing the show to continue remains utterly unacceptable to Una.

Thus, immediately helping Una minimizes the maximum degree of unacceptability to any individual. By Minimax, we must help Una. Similar reasoning applies to any situation where one person has much more at stake than the others.

§3.3: Quantity Versus Quality

One distinctive implication of the theory emerges in choices between helping a few more people and helping people a bit more.

Suppose that the local schools are struggling to pay for new textbooks. The current books are worn and a decade out of date. We can fund new textbooks at Eastview High School or Southview High School. We cannot afford to fund both, nor would it make sense to provide textbooks for only some students at each school. The students at both schools are similar and would benefit similarly from updated materials.

The schools have different numbers of students. Eastview High School has a thousand students, whereas Southview has just nine hundred. So if we bought textbooks for Eastview, we could only afford the previous edition, whereas if we bought textbooks for Southview, we could buy the latest edition. We estimate that the potential benefit to each Eastview student from an upgrade to the previous edition gives them about ninety percent of the educational benefit that each Southview student could get from an upgrade to the latest edition. So our

options are to either purchase the previous edition for a thousand students at Eastview, or purchase the latest edition for nine hundred students at Southview. How does the theory apply?

Imagine purchasing the books for Eastview. This benefits a thousand students at Eastview while leaving nine hundred students at Southview with the old textbooks. The best case for this involves asking each student at Southview to consider the perspective of a corresponding student at Eastview. This leads each student at Southview to assess our action more favorably. Each may still find the action unacceptable, since they stood to gain even more from upgrading to the latest editions than their counterpart Eastview student gains from upgrading to the previous edition, but as the difference is relatively small, each Southview student finds our action only mildly unacceptable.

Imagine purchasing the books for Southview. This benefits nine hundred students at Southview while leaving a thousand students at Eastville with the old textbooks. The best case for this involves asking nine hundred of the students at Eastville to each consider the perspective of a corresponding student at Southville. This mollifies those students. But to the remaining hundred Eastville students, we have nothing to say. They can continue to find the action highly unacceptable, given how much they stood to gain from the updated textbooks.

Thus, purchasing the books for Eastview minimizes the maximum degree of unacceptability to any individual. By Minimax, we must buy the books for Eastview.

This conclusion departs from typical aggregative moral theories, which allow improvements in quantity and quality to trade off smoothly, thereby permitting one to buy books for Eastview (where eleven percent more people stand to benefit) or Southview (where each person stands to benefit eleven percent more). This also departs from 'partially

aggregative' theories, which maintain that one should aggregate when people each have sufficiently similar goods at stake, thereby yielding the same implications as aggregative theories in cases like this one.[16]

Strikingly, the above reasoning applies even when only one more student attends Eastview than Southview. Whenever the stakes for everyone are broadly similar, the theory directs us to benefit as many people as possible. I discuss this feature further in §4.

§3.4: How the Theory Works in General

When working through these examples, we considered only the pattern of internalization where those that favor one action each internalize the perspective of someone that favors the other action, at least so far as this was possible. We can generalize this structure.

Given any two options, let Top-Down be the pattern of internalization that asks the person with the most to gain from one option to consider the perspective of the person with the most to gain from the other option, that asks the person with the second-most to gain from one option to consider the perspective of the person with the second-most to gain from the other option, and so on. Top-Down asks those with more at stake to consider each others' opposed perspectives.

It turns out that Top-Down patterns of internalization are always one of the best case eligible patterns of internalization. No matter which options we compare, the best case for the former option relative to the latter is the one Top-Down prescribes. As a result, Minimax implies that we need only consider the Top-Down patterns of internalization.[17]

---

[16] For an overview, see Horton 2021.

[17] See Fact 2 of the appendix.

So when we think through the consequences of the theory, we do not need to actually consider every eligible pattern of internalization. We do not need to think about what happens if no one internalizes anyone's perspective, or only eight students internalize anyone's perspective, and so on. It suffices to think about the Top-Down pattern, where those with more to gain from one action each consider the perspective of someone with more to gain from the other action, while those with less to gain from an action each consider the perspective of those with less to gain from the alternative. When an action minimizes the maximum degree of unjustifiability to anyone given this pattern, it is permissible.

§4: Objection and Reply

The most distinctive implication of the theory is that when people's interests are broadly similar, we must help as many people as possible. This effaces distinctions between cases that intuitively matter and that many competing theories draw. For example, when considering the following:

- You can give one million students the latest edition textbook or another one million and one students the previous edition
- You can give one student the latest edition textbook or another one million students the previous edition

the theory implies that one must distribute the previous edition in both choices. Intuitively, though, we should distribute the latest edition in the first choice and the previous edition in the second choice.

The reason for this implication, as always, comes down to Minimax. In both choices, however we ask people to consider others' perspectives, there will be a student who could benefit from an upgrade to the previous edition and who does not internalize anyone else's perspective. Distributing the latest edition, in either choice, remains fairly unacceptable to them. And in both choices, we can ask everyone who would benefit from the latest edition textbook to each internalize the perspective of someone who would benefit from the previous edition textbook. Distributing the previous edition, in either choice, becomes only mildly unacceptable to each of them. Thus, distributing the previous edition minimizes the maximum degree of unacceptability to anyone in both choices. By Minimax, this is what we must do.

This example may not seem so objectionable, perhaps because not much depends on the textbook one receives. But similar reasoning applies to choices with larger stakes:

- You can fund K-12 education for a million students or fund K-5 education for another million and one students
- You can fund K-12 education for one student or fund K-5 education for another million students

It seems clear that one should fund K-12 education in the first choice but not the second. But the theory implies that one may fund K-12 education in the first choice iff one may fund K-12 education in the second choice.

This is because, as before, however we ask people to consider others' perspectives, there will be a child who could receive a K-5 education and who does not internalize anyone else's perspective. Funding the K-12 programs, in either choice, would be highly unacceptable to this child. And in both choices, we could ask everyone who would benefit from the K-12 education to each internalize the perspective of someone who would benefit from the K-5 education. Since they have so much more to gain, though, funding the K-5 programs may remain highly unacceptable to each of them. The key question is how these degrees of unacceptability compare. If the former is weakly greater, then one may fund the K-5 programs in both choices. If the latter is strictly greater, then one must not fund the K-5 programs in either choice. In either case, one may fund K-12 education in the first choice iff one may fund K-12 education in the second choice.

How should a proponent of the theory react to these implications?

One option is to revise the theory to accomodate these judgments. Perhaps individuals can be asked to internalize the perspectives of multiple others or to consider the sum of others'

wellbeing after all. Perhaps this can be maintained while also insisting that there is a set of cases, such as the theatre case, where these mechanisms do not apply. The idea might be that when individual fates differ dramatically, we must consider individuals as individuals, but when individuals face similar stakes, we are allowed to group them together in our moral thought.[18] The revised theory would still offer a distinctive account of what it means to treat individuals as individuals while, perhaps, accomodating our intuitive reactions to cases.

This is a live option, but for perhaps idiosyncratic reasons, it does not appeal to me. A view where individuals differ fundamentally from groups, so that one simply cannot internalize the perspectives of multiple others or sum the wellbeing of multiple others, somehow seems to me a genuine possibility. Things really could be that way. I struggle to feel similarly about a view that allows for these things while circumscribing their moral significance in accordance with our intuitions about cases. This feels to me like special pleading, an attempt to save the theory from devastating counter-examples by the least disruptive means available. As it stands, the theory's principles seem to me to form a unified package, emerging naturally from the nature of internalization, and I am loath to give up this feature.

Another option is to reject the theory on the basis of these examples. That may be the right conclusion to draw in the end, but there is another option worth considering first.

My preferred option is to consider how making these cases more realistic might alter the prescriptions of the theory. The above examples make several unrealistic assumptions, such as that (i) the agent has only two extreme options, and (ii) the agent makes a single

[18] This justification, which sounds in ideas of disrespect and fungibility, differs somewhat from Voorhoeve 2014: 70 – 75's appeal to personal prerogatives and the obligation to withdraw sufficiently weak claims. It also differs from Henning 2023: 771's appeal to autonomy and Zhang 2024: 490 – 491's appeal to the impropriety of counting certain considerations against one another. These alternatives may provide more plausible in the end, though.

decision. But resources can typically be allocated in flexible ways, so an agent who can fund many K-12 or many K-5 scholarships can typically also fund some of each; it is also rare for each available option to benefit a disjoint group. Moreover, agents rarely make decisions with isolated moral significance; rather, a suite of preliminary, ongoing, and subsequent decisions shape the outcome and significance of particular decisions. These decisions, in turn, are shaped by the decisions of others, so that one should sometimes think of one's decision as part of a larger suite of decisions whose moral significance must be assessed jointly. Excepting egregious cases, it is not the individual tax collector whose actions come in for moral assessment but rather the system of taxation as a whole.

Dropping these unrealistic features can lead the theory to make more attractive prescriptions. For example, suppose the potential beneficiaries of the educational funding overlap:

- You can fund K-12 education for a million students or fund K-5 education for those same million students plus an additional student
- You can fund K-12 education for one student or fund K-5 education for that same student plus another million students

The theory can then require that you fund K-12 education in the first choice and K-5 education in the second choice.

This is because, in the first choice, we can now ask the additional student to consider the importance of providing K-12, rather than K-5, education to someone else. This can make choosing to fund K-12 education acceptable to them (or, at least, less unacceptable than failing to provide someone with K-12, rather than K-5, education would be to someone else). Thus, funding K-12 education minimizes the maximum degree of unacceptability to anyone, so by

Maximin, we must do so. In the second choice, we can ask the student who stands to gain a K-12 education to consider the importance of providing K-5 education to someone else. This can make choosing the latter acceptable to them (or, at least, less unacceptable than failing to provide someone with K-5 education would be to someone else). Thus, funding K-5 education minimizes the maximum degree of unacceptability to anyone, so by Maximin, we must do so.

It would take a separate paper to assess, more generally, how the theory fares when we make these cases more realistic.[19] But these challenges to the theory rely on sharp contrasts between specific options, so broadening those option sets in realistic ways may well lead the theory to make more familiar prescriptions.[20]

Even if the theory makes familiar prescriptions in realistic cases, one might think: So what? True moral theories apply to all cases, so an erroneous prescription in any case refutes the theory. We can safely reject this theory given its odd prescriptions in the above cases, regardless of how it performs elsewhere.

But I doubt the reliability of our moral intuitions in highly unrealistic cases.[21] Another possibility is that our intuitions are deeply shaped by background presumptions about how the world typically works and that they emerge from a training process suffused with idiosyncracies of the local culture. The judgments we have about the cases above may reflect the ineliminable influence of the fact that we actually live in a world that raises constant,

---

[19] Chapter 3 of my dissertation uses simulations to argue that the theory yields sensible prescriptions in cases where options are drawn from uniform and normal distributions, as seems realistic.

[20] Prominent alternatives, such as contractualism, face similar structural problems that proponents similarly address by focusing on general policies in realistic cases. Uncertainty about what the general policy will do (e.g. James 2012: 282 – 283; Wallace 2019: 215 – 221) or intrapersonal aggregating over the different ways the policy will affect individuals over the course of their whole lives (Scanlon 1998: 237; Tadros 2019: 201 – 202) lead contractualism to make intuitively sensible conclusions, or so these authors argue. Similar remarks may apply to the theory of this paper.

[21] This parallels standard defenses to counterintuitive implications of act utilitarianism (e.g. Smart 1973: 72 –73).

complex tradeoffs of different kinds, across different domains, across different people; and that we coordinate in limited but remarkable and ever increasing ways across such choices. The simple principles of the theory, plus the complex empirical realities we face, may be together needed to explain why certain actions are appropriate. And when those complex empirical realities do not obtain, such actions may not be appropriate even though we struggle to recognize this.

Conclusion

We can all be asked to take up another person's perspective, to take seriously what they have at stake, and to modify our assessment of various actions accordingly. This process, which we term internalization, suggests three principles: No Multitudes, No Multiplicity, and Minimax. The theory of this paper conjoins these principles, and this theory matches our judgments about the transit and theatre cases with which we began. The theory also makes odd prescriptions in certain cases, but perhaps the oddity can be attributed to odd features of the cases rather than the theory.

One final attraction of the theory concerns the kind of justification it makes available to those who lose out. If someone acts according to this theory and things nonetheless turn out poorly for me, I can generally console myself with the knowledge that there actually is someone with something similar at stake whose interests are satisfied, although mine are not. I lose out, but there is someone who correspondingly gains. When social institutions accord with the theory, those who lose out can know that their interests are never sacrificed to the maw of impersonal utility; rather, they are generally sacrificed only to satisfy the similar or weightier interests of someone else. The mechanisms involved may be incredibly complex and there may be no way in practice to identify their counterpart, but there is actually someone who benefited on the other side. Such justifications may be more convincing than others.

Appendix

We establish the number of eligible patterns of internalization (Fact 1), and we show that in any choice, it suffices to consider only the Top-Down patterns of internalization (Fact 2).

Definitions and Assumptions

$I = \{1,2,\ldots,n\}$ is the set of individuals.

$O = \{\varphi, \ldots, \psi\}$ is the set of options.

$\sigma(\phi, \psi, i) : O^2 \times I \to \mathbb{R}$ represents how choosing ϕ over ψ matters for i. We abbreviate this as $\sigma_i^{\phi,\psi}$.

We assume:

1. $\sigma_i^{\phi,\psi}$ is positive when i's interests favor ϕ over ψ, negative when they ψ over ϕ, and otherwise zero.
2. $|\sigma_i^{\phi,\psi}|$ increases as the choice between ϕ and ψ matters more for i.

$P \subseteq I \times I$ is a pattern of internalization. iPj means i internalizes j's perspective.

A binary relation R is functional iff $aRb \wedge aRc \implies b = c$.

A binary relation R is injective iff $aRc \wedge bRc \implies a = b$.

No Multitudes: a pattern of internalization is eligible only if it is functional.

No Multiplicity: a pattern of internalization is eligible only if it is injective.

Sufficiency: if a pattern of internalization is injective and functional, then it is eligible.[22]

$\mathscr{E}$ is the set of all eligible patterns of internalization.

[22] This includes patterns of internalization where someone internalizes their own perspective. These are not obviously coherent, so one may want to rule them out. One easy fix is to require that P is irreflexive while amending the definition of Top-Down and the formula in Fact 1 accordingly. Nothing much depends on this.

Given options ϕ and ψ, define Top-Down(ϕ, ψ) as follows:

$$p_1 = \min \arg \max_{i \in I} \sigma_i^{\phi,\psi} \text{ and } p_k = \min \arg \max_{i \in I \setminus \{p_1,\ldots,p_{k-1}\}} \sigma_i^{\phi,\psi}$$

$$q_1 = \min \arg \max_{i \in I} \sigma_i^{\psi,\phi} \text{ and } q_k = \min \arg \max_{i \in I \setminus \{q_1,\ldots,q_{k-1}\}} \sigma_i^{\psi,\phi}$$

$$\text{Top-Down}(\phi, \psi) = \{(p_k, q_k) \mid 1 \leq k \leq n\}$$

We use $TD^{\phi,\psi}$ to abbreviate Top-Down(ϕ, ψ).

$u(i, j, \varphi, \psi) : I^2 \times O^2 \to \mathbb{R}$ represents the degree of unacceptability to i of choosing ϕ over ψ even after i internalizes j's perspective.

We assume:

The unacceptability of an action increases as the alternative improves:

$u(i, j, \varphi, \psi)$ weakly increases as $\sigma_i^{\psi,\phi}$ increases

The unacceptability of an action increases as it benefits the other less:

$u(i, j, \varphi, \psi)$ weakly increases as $\sigma_j^{\psi,\phi}$ decreases

Internalizing no one's perspective leaves one's outlook unchanged:

$$u(i, \varnothing, \varphi, \psi) = \sigma_i^{\phi,\psi}$$

Internalizing the perspective of someone with nothing at stake is like considering no one:

$$\text{If } \sigma_j^{\phi,\psi} = \sigma_j^{\psi,\phi} = 0, \text{ then } u(i, j, \varphi, \psi) = u(i, \varnothing, \varphi, \psi)$$

Internalizing the perspective of someone who favors the same action to an even greater degree leads one to adopt their perspective:

$$\text{If } \sigma_j^{\phi,\psi} > \sigma_i^{\phi,\psi}, \text{ then } u(i, j, \varphi, \psi) = \sigma_j^{\phi,\psi}$$

Minimax: one may choose all and only the members of

$$\arg \min_{\phi \in O} \max_{\psi \in O} \min_{P \in \mathscr{E}} \max_{i \in I, [\exists k: iPk \rightarrow iPj], [\neg \exists k: iPk \rightarrow j = \varnothing]} u(i, j, \phi, \psi)$$

Fact 1

Given n people, the number of eligible patterns of internalization is

$$\sum_{k=0}^{n} \binom{n}{k}^2 k!$$

Proof

Let $I = \{1,2,...,n\}$. The eligible patterns of internalization are all binary relations $P \subset I^2$ that are injective and functional. Each pattern of internalization can be seen as a directed graph with n nodes. We can categorize these graphs by their number of arrows. The eligible patterns of internalization are exactly those directed graphs where the number of arrows equals the number of distinct heads of the arrows and the number of distinct tails of the arrows.

Given k arrows and n nodes, there are $\binom{n}{k}$ ways to choose k distinct heads for the arrows. Similarly, there are $\binom{n}{k}$ ways to choose k distinct tails for the arrows. And there are $k!$ ways to associate the given tails with the given heads. Thus, there are $\binom{n}{k}^2 k!$ directed graphs with k arrows, k distinct heads, and k distinct tails that can be made from n nodes. Summing over all possible k yields our formula.

Fact 2

Assume No Multitudes, No Multiplicity, Sufficiency, and Minimax. Then, one may choose all and only the members of:

$$\arg\min_{\phi\in O}\max_{\psi\in O}\max_{i\in I,[\exists k:iTD^{\phi,\psi}k\to iTD^{\phi,\psi}j],[\neg\exists k:iTD^{\phi,\psi}k\to j=\varnothing]} u(i,j,\phi,\psi)$$

Proof

By Minimax, one may choose all and only the members of:

$$\arg\min_{\phi\in O}\max_{\psi\in O}\min_{P\in\mathscr{E}}\max_{i\in I,[\exists k:iPk\to iPj],[\neg\exists k:iPk\to j=\varnothing]} u(i,j,\phi,\psi)$$

It suffices to show that for all ϕ, ψ:

$$\min_{P\in\mathscr{E}}\max_{i\in I,[\exists k:iPk\to iPj],[\neg\exists k:iPk\to j=\varnothing]} u(i,j,\phi,\psi) = \max_{i\in I,[\exists k:iTD^{\phi,\psi}k\to iTD^{\phi,\psi}j],[\neg\exists k:iTD^{\phi,\psi}k\to j=\varnothing]} u(i,j,\phi,\psi)$$

By Sufficiency, $TD^{\phi,\psi}\in\mathscr{E}$. Thus:

$$\min_{P\in\mathscr{E}}\max_{i\in I,[\exists k:iPk\to iPj],[\neg\exists k:iPk\to j=\varnothing]} u(i,j,\phi,\psi) \leq \max_{i\in I,[\exists k:iTD^{\phi,\psi}k\to iTD^{\phi,\psi}j],[\neg\exists k:iTD^{\phi,\psi}k\to j=\varnothing]} u(i,j,\phi,\psi)$$

So it suffices to show:

$$\min_{P\in\mathscr{E}}\max_{i\in I,[\exists k:iPk\to iPj],[\neg\exists k:iPk\to j=\varnothing]} u(i,j,\phi,\psi) \geq \max_{i\in I,[\exists k:iTD^{\phi,\psi}k\to iTD^{\phi,\psi}j],[\neg\exists k:iTD^{\phi,\psi}k\to j=\varnothing]} u(i,j,\phi,\psi)$$

Let i* be someone to whom choosing ϕ rather than ψ is maximally unjustifiable under $TD^{\phi,\psi}$.

In other words,

$$u(i^*,k^*,\phi,\psi) = \max_{i\in I,[\exists k:iTD^{\phi,\psi}k\to iTD^{\phi,\psi}j],[\neg\exists k:iTD^{\phi,\psi}k\to j=\varnothing]} u(i,j,\phi,\psi)$$

where k* = ∅ if $\neg\exists k : i^*TD^{\phi,\psi}k$, and otherwise $i^*TD^{\phi,\psi}k^*$.

Let $A = \{i \in I : \sigma_i^{\phi,\psi} \geq \sigma_{i^*}^{\phi,\psi}\}$. If $\text{k}^* \neq \emptyset$, let $B = \{j \in I : \sigma_j^{\psi,\phi} > \sigma_{k^*}^{\psi,\phi}\}$. Otherwise, let $B = \{j \in I : \sigma_j^{\psi,\phi} > 0\}$. By definition of $TD^{\phi,\psi}$, #A > #B. Thus by No Multitudes and No Multiplicity, on any eligible pattern, some member of A either internalizes no one's perspective or internalizes the perspective of exactly one person who is not a member of B. So for any $P \in \mathscr{E}$ there exists $a \in A$ such that $\neg\exists k : aPk$ or $\exists! c : aPc$ where $c \notin B$. Thus:

$$\max_{i\in I,[\exists k:iPk\rightarrow iPj],[\neg\exists k:iPk\rightarrow j=\emptyset]} u(i,j,\phi,\psi) \geq u(a,c,\phi,\psi) \geq u(i^*,c,\phi,\psi) \geq u(i^*,k^*,\phi,\psi)$$

where the latter two inequalities follow from the above assumptions about u(·). Thus:

$$\max_{i\in I,[\exists k:iPk\rightarrow iPj],[\neg\exists k:iPk\rightarrow j=\emptyset]} u(i,j,\phi,\psi) \geq \max_{i\in I,[\exists k:iTD^{\phi,\psi}k\rightarrow iTD^{\phi,\psi}j],[\neg\exists k:iTD^{\phi,\psi}k\rightarrow j=\emptyset]} u(i,j,\phi,\psi)$$

Since $P \in \mathscr{E}$ was arbitrary,

$$\min_{P\in\mathscr{E}} \max_{i\in I,[\exists k:iPk\rightarrow iPj],[\neg\exists k:iPk\rightarrow j=\emptyset]} u(i,j,\phi,\psi) \geq \max_{i\in I,[\exists k:iTD^{\phi,\psi}k\rightarrow iTD^{\phi,\psi}j],[\neg\exists k:iTD^{\phi,\psi}k\rightarrow j=\emptyset]} u(i,j,\phi,\psi)$$

as desired.

Bibliography

Bailey, Olivia (2021). Empathy with vicious perspectives? A puzzle about the moral limits of empathetic imagination. *Synthese* 199 (3-4):9621-9647.

Bailey, Olivia (2022). Empathy and the Value of Humane Understanding. *Philosophy and Phenomenological Research* 104 (1):50-65.

van Gils, Aart & Tomlin, Patrick (2020). Relevance Rides Again?. *Oxford Studies in Political Philosophy Volume 6.* Edited by David Sobel, Peter Vallentyne & Steven Wall. Oxford University Press.

Goldie, Peter (2000). *The Emotions: A Philosophical Exploration*. Oxford University Press.

Gustafsson, Johan E. (2021). Utilitarianism without Moral Aggregation. *Canadian Journal of Philosophy* 51 (4):256-269.

Hart, James (2022). Tie-breaks and Two Types of Relevance. *Ethical Theory and Moral Practice* 25 (2):1-20.

Hart, James (2024). Limited Aggregation's Non-Fatal Non-Dilemma. *Australasian Journal of Philosophy*.

Henning, Tim (2023). Numbers without aggregation. *Noûs* (3):755-777.

Hirose, Iwao (2001). Saving the greater number without combining claims. *Analysis* 61 (4):341-342.

Horton, Joe (2017). Aggregation, Complaints, and Risk. *Philosophy and Public Affairs* 45 (1):54-81.

Horton, Joe (2021). Partial aggregation in ethics. *Philosophy Compass* 16 (3):1-12.

James, Aaron (2012). Contractualism's (Not So) Slippery Slope. *Legal Theory* 18 (3):263-292.

Kamm, Frances Myrna (1993). *Morality, Mortality Volume I.* Oxford University Press.

Kamm, Frances Myrna (2006). *Intricate Ethics*. Oxford University Press.

Kumar, Rahul (2011). Contractualism on the Shoal of Aggregation. *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon.* Edited by R. Jay Wallace, Rahul Kumar, and Samuel Freeman. Oxford University Press.

Mann, Kirsten (2021). Relevance and Nonbinary Choices. *Ethics* 132 (2):382-413.

Raz, Joseph (2003). Numbers, With and Without Contractualism. *Ratio* 16 (4):346–367.

Reibetanz, Sophia (1998). Contractualism and aggregation. *Ethics* 108 (2):296-311.

Rüger, Korbinian (2020). Aggregation with Constraints. *Utilitas* 32 (4):454-471.

Scanlon, Thomas. (2021). Contractualism and Justification. *Reason, Justification, and Contractualism: Themes from Scanlon*. Edited by Markus Stepanians & Michael Frauchiger. De Gruyter.

Scanlon, Thomas. (1998). *What We Owe to Each Other*. Harvard University Press.

Smart, J. J. C. & Williams, Bernard (1973). *Utilitarianism: For and Against*. Cambridge University Press. Edited by Bernard Williams.

Steuwer, Bastian (2021). Aggregation, Balancing, and Respect for the Claims of Individuals. *Utilitas* 33 (1):17-34.

Tadros, Victor (2019). Localized Restricted Aggregation. *Oxford Studies in Political Philosophy Volume 5*. Edited by David Sobel, Peter Vallentyne and Steven Wall. Oxford University Press.

Taurek, John (1977). Should the Numbers Count? *Philosophy and Public Affairs* 6 (4):293-316.

Taurek, John (2021). Reply to Parfit's 'Innumerate Ethics.' *Principles and Persons: The Legacy of Derek Parfit*. Edited by Jeff McMahan and others. Oxford University Press.

Voorhoeve, Alex (2014). How Should We Aggregate Competing Claims. *Ethics* 125 (1):64-87.

Voorhoeve, Alex (2017). Why One Should Count only Claims with which One Can

Sympathize. *Public Health Ethics* 10:2:148-56.

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

Zhang, Erik (2024). Individualist Theories and Interpersonal Aggregation. *Ethics* 134 (4):479-511.