



ARTICLES

# Reciprocity, Inequality, and Unsuccessful Rescues

Romy Eskens 

Department of Philosophy and Religious Studies, Utrecht University, Utrecht, The Netherlands  
Email: [r.s.eskens@uu.nl](mailto:r.s.eskens@uu.nl)

## Abstract

Forced choices between rescuing imperilled persons are subject to a presumption of equality. Unless we can point to a morally relevant difference between these persons' imperillments, each should get an equal chance of rescue. Sometimes, this presumption is overturned. For example, when one imperilled person has wrongfully caused the forced choice, most think that *this* person (rather than an innocent person) should bear the harm. The converse scenario, in which a forced choice resulted from the supererogatory action of one of the imperilled people, has received little attention in distributive ethics. I argue that, sometimes, we need not offer equal chances in these cases either. When the supererogatory act places the initially imperilled person under a reciprocal duty to bear risks for the supererogatory agent's sake in the forced choice, we may fulfil this duty for them if they are unable to do it themselves, by favouring the supererogatory agent.

**Keywords:** distributive ethics; ethics of rescuing; beneficence; reciprocity; gratitude

## 1. Introduction

When a rescuer can save only one of two equally imperilled people, whom it is equally difficult to save, they should ordinarily give each person an equal chance of rescue. However, when one imperilled person has wrongfully caused the forced choice, most think that *this* person (rather than an innocent person) should bear the harm. This claim underpins, for example, Jeff McMahan's influential account of liability to defensive harm.<sup>1</sup> The converse scenario, in which a forced choice results from the supererogatory action of one of the imperilled persons, has received very little attention in distributive ethics.<sup>2</sup> In this paper, I investigate one such scenario, in which someone's unsuccessful supererogatory attempt to save an imperilled person creates a forced choice between their lives. I argue that, sometimes, we need not be impartial in these

<sup>1</sup>See, for instance, McMahan (2005).

<sup>2</sup>Driver (1997) discusses supererogatory beneficial interventions that go awry. However, Driver does not consider cases in which such interventions cause a forced choice between the would-be beneficiary and the would-be benefactor. There are brief discussions of forced choices caused by obligatory (but not supererogatory) actions in McMahan (2009: 165), Otsuka (2009: 383–83), and Tadros (2011: 232–35).

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

cases either. Sometimes, we may (or even should) give the unsuccessful rescuer a greater chance of rescue than the initially imperilled person.

Forced choices caused by unsuccessful supererogatory rescue attempts are not of merely theoretical interest. They are also of considerable practical importance. Professional rescuers, such as firefighters and combatants, must sometimes choose between the lives of unsuccessful rescuers and those of other imperilled people. Moreover, it is often the *risks* of harm, rather than the harms themselves, which need distributing. For example, decision-making about the strategy of a justified humanitarian intervention involves distributing risks of harm between the intervening combatants and the civilians whom these combatants are trying to save. When a military commander orders an aircrew to bomb targets from high altitudes, for instance, the crew is less likely to be shot down than if they were to bomb from lower altitudes, but civilians are more likely to be collaterally killed compared to that alternative.<sup>3</sup>

I proceed as follows. In Section 2, I explain the presumption of equality that ordinarily obtains in forced choices, and present an example of a forced choice caused by supererogation in which this presumption is plausibly overturned. In Section 3, I argue that the considerations typically invoked to justify unequal distributions in forced choices cannot justify such a distribution in this example. In Sections 4 and 5, I suggest filling the gap with a reciprocity-based view: when the initially imperilled person has a duty to reciprocate the unsuccessful rescuer's attempt to save them, by bearing risks for their sake in the forced choice, the decision-maker may (or even should) see to it that this duty is fulfilled if the initially imperilled person cannot fulfil it themselves, thus favouring the unsuccessful rescuer. Section 6 concludes.

## 2. Forced choices and the presumption of equality

Forced choices between rescuing imperilled persons are subject to a presumption of equality. Unless we can point to some morally significant difference between the imperilled persons with respect to their imperilments, each person should get an equal chance of being saved.<sup>4</sup> In many forced choices, no such difference exists. This is true, for instance, in *Innocent Trolley*:

### *Innocent Trolley*

Due to a natural misfortune, Anne and Beth are each stuck on a track on which a trolley is travelling. If no one saves them, they will certainly die. Claire can pull either Anne or Beth to safety, but not both.

There is no morally significant difference between Anne and Beth with respect to their imperilments. If we assume that all else is also equal between them, Claire should give Anne and Beth equal chances of rescue. Claire may not simply save one rather than the other, nor may she give one a greater chance of rescue than the other.

But now consider:

### *Evil Trolley*

Beth culpably ties Anne to a track on which a trolley is travelling, in order to unjustly kill Anne. When fleeing the scene, Beth gets stuck on another track, on

<sup>3</sup>For moral evaluation of such a strategy, see Øverland (2011).

<sup>4</sup>For this claim, see, for instance, Frowe (2021: 70) and Tadros (2011: 203, 211).

which another trolley is travelling. If no one saves Anne and Beth, they will certainly die. Claire can pull either Anne or Beth to safety, but not both.

In this case, there are morally significant differences between Anne's and Beth's imperilments: (i) Beth wrongfully caused Anne's imperilment, while Anne did not wrongfully cause Beth's imperilment; (ii) Beth is at fault for her own imperilment, while Anne is not at fault for her own imperilment. These differences plausibly overturn the presumption of equality that ordinarily obtains in forced choices. Rather than give Anne and Beth equal chances of being saved, Claire ought to save Anne rather than Beth.

Next, consider the following, more elaborate example, which is the focal point of this paper:

### *Heroic Trolley*

Due to a natural misfortune, Anne is stuck on a track on which a trolley is travelling. If no one saves her, she will certainly die. Claire is running towards the levers that will stop the trolley, but she has only a 0.3 chance of reaching them in time to save Anne. Beth knows that she has a 0.7 chance of pulling Anne to safety, but she also knows that if she fails, she will get stuck on a different track on which another trolley is travelling.<sup>5</sup> Beth attempts the rescue, but is unsuccessful: she ends up stuck on the other track. Claire reaches the levers in time to certainly save either Anne or Beth, but she cannot save both. Anne, Beth, and Claire do not know each other, and cannot communicate. Because they cannot communicate, Anne could not consent to Beth's attempt, and Claire does not know Anne's or Beth's preferences about whom to save in the forced choice.

This example resembles cases in which civilians undertake supererogatory attempts to rescue others, but their attempts fail and a forced choice ensues. It also resembles real-world scenarios involving professional rescuers. For example, a firefighter may be forced to choose between the life of a civilian and that of a fellow firefighter. One might think that examples involving professional rescuers are relevantly different from *Heroic Trolley* because, unlike civilian rescuers, these rescuers are *required* to save others at great risk to themselves. However, there are limits even to the risk that professional rescuers can be required to incur to save others. When their attempt to save someone exceeds this limit and creates a forced choice, the resulting scenario is again similar to our example.

In *Heroic Trolley*, Beth attempts to save Anne, but ends up making things worse for her: had Beth not intervened, only Anne would need saving, and Claire would have been sure to save her. Still, I think that Beth acts permissibly in attempting the rescue.

The main consideration in support of this judgement is based on hypothetical consent. At the time of decision-making, the attempt increases Anne's prospect of survival from 0.3 to 0.7. This increase seems substantial enough to make it irrational for Anne to refuse to consent to the attempt and, hence, likely that she would consent if she were able to.<sup>6</sup> In the absence of Anne's actual consent (or dissent) or knowledge about

<sup>5</sup>It is important to introduce specific probabilities already here (rather than simply saying "good chance" and "small chance"). I want to focus on unsuccessful rescue attempts that are permissible and supererogatory, and whether they are depends on the balance of probabilities. See the next page and Section 3.1 for elaboration.

<sup>6</sup>Of course, Anne's refusal would not be irrational if she had access to all the facts – most importantly, the fact that Claire will reach the levers in time to save her. But Anne does not have access to these facts at

whether the attempt will succeed, Beth seems permitted to proceed. (In Section 3.1, I say more in support of this claim and in response to possible objections.)

Beth is permitted to proceed, but probably not required. Recall that if Beth's attempt fails, Beth will get stuck on the other track. The attempt has a 0.7 chance of succeeding, so the risk of Beth's getting stuck on the other track is 0.3. If Claire gives Anne and Beth equal chances of being saved should a forced choice arise, Beth exposes herself to a 0.15 risk of dying by attempting the rescue.<sup>7</sup> Even if Claire gives Beth a greater chance of rescue than Anne (as I argue she should), Beth still takes on a serious risk. For instance, if Claire gives Beth a chance of 0.7, the risk of dying that Beth subjects herself to in attempting to save Anne is 0.09.<sup>8</sup> Plausibly, morality does not demand that we bear risks of these magnitudes to save a stranger's life. Beth's attempt is supererogatory.<sup>9</sup>

This paper discusses only forced choices of this kind, caused by supererogation. It does not address forced choices caused by morally required rescue attempts, because the introduction of requirements adds complexities that I cannot take up here. However, I elsewhere argue that we can incur reciprocal duties in response to required (and even rights-conforming) acts of benefitting.<sup>10</sup> If I am right about this, the reciprocity-based argument that I develop in this paper likely extends to unsuccessful required (and even rights-conforming) rescues.

Let us now turn to our main question. Is there a morally significant difference between Anne's and Beth's imperilments in *Heroic Trolley* that overturns the presumption of equality? If so, is the permission to treat Anne and Beth unequally in Anne's or Beth's favour? I suspect that most readers will answer the first question in the affirmative: there is a relevant asymmetry between Anne's and Beth's imperilments that bears on what Claire is permitted to do in the forced choice.<sup>11</sup> In this paper, I grant that claim in order to focus on the second question: whom does the asymmetry favour, and why? My view is that Claire ought to favour Beth because this fulfils Anne's reciprocal duty towards Beth, which Anne is unable to fulfil herself. However, before presenting and defending that view, I need to explain why the asymmetries typically appealed to in order to justify distributive inequality in forced choices are ill-suited for justifying such inequality in *Heroic Trolley*.

### 3. Distributive inequality

Readers familiar with distributive ethics might want to appeal to an asymmetry in wrongdoing, option luck, or autonomous preferences to justify favouring Anne, or an asymmetry in beneficiary liability to justify favouring Beth. I consider the resulting views in turn. In each case, I grant that the asymmetry in question sometimes justifies

---

the time she would be dissenting to Beth's attempting the rescue. At that time, Anne knows only that Beth's attempting the rescue increases her chance of survival from 0.3 to 0.7. Given this knowledge, dissenting is irrational.

<sup>7</sup>Multiply the 0.3 risk of getting stuck by the 0.5 risk of not being saved by Claire should a forced choice arise.

<sup>8</sup>Multiply the 0.3 risk of getting stuck by the 0.3 risk of not being saved by Claire should a forced choice arise.

<sup>9</sup>Those who disagree can change the probabilities as they see fit.

<sup>10</sup>Eskens (manuscript).

<sup>11</sup>Not everyone shares this intuition. Those who do not can read the reciprocity-based argument of this paper as an attempt to convince them that, contrary to their intuitions, there *is* an asymmetry that bears on what Claire is permitted to do (rather than as an attempt to explain and justify a prior intuition).

prioritising one imperilled person relative to another. However, I argue that each of the asymmetries either does not arise in the case that I want to solve, or lacks moral relevance in that case.

### 3.1. *Wrongdoing*

I start with asymmetrical wrongdoing. Beth's rescue attempt increases Anne's chance of survival from 0.3 to 0.7. Above, I suggested that this increase is substantial enough to make it irrational for Anne to dissent to the attempt and likely that she would consent if she were able to. I also suggested that, in the absence of Anne's actual consent (or dissent) or knowledge about whether the attempt will succeed, these facts make it permissible for Beth to proceed.

Some might dispute this. Although Beth increases Anne's overall prospect of survival by attempting to save her, she also imposes a new risk on Anne: the risk of not being saved by Claire should a forced choice arise. One might think that the imposition of this new risk on Anne requires Anne's actual consent to be permissible, and this consent is lacking. If this is right, Beth acts wrongly in attempting to save Anne. And if Beth acts wrongly in attempting to save Anne, Claire should prioritise saving Anne, just as in *Evil Trolley*.

I doubt that an additional justificatory burden attaches to Beth's imposing the new risk on Anne. The imposition is a means of increasing Anne's overall chance of rescue, and it is rational for Anne to consent to this increase even if the risk of not being saved in the forced choice is factored in. Since it is the irrationality of Anne's dissent and the corresponding likelihood that Anne would consent to Beth's intervening if she were able to that makes it permissible for Beth to attempt the rescue in the absence of Anne's consent, these same factors should also make it permissible to impose the new risk included in the attempt.

Here is a different argument for Beth's attempt being impermissible: it has bad consequences. The attempt is of course expected to benefit Anne, but it actually makes her worse off. Given this, Beth should not attempt the rescue; attempting it is wrong.<sup>12</sup>

This argument assumes a fact-relative understanding of wrongdoing, according to which an act's wrongness depends on whether the agent would be permitted to perform the act if they knew all the facts.<sup>13</sup> By contrast, I have so far been assuming an evidence-relative understanding, which lets an act's wrongness depend on whether the agent is permitted to perform it in light of the available evidence at the time of decision-making.<sup>14</sup> There is extensive debate in normative ethics about whether fact-relative or evidence-relative wrongdoing is primary.<sup>15</sup> This is not the place to delve deeply into this debate. However, I want to briefly mention a famous example, due to Frank Jackson (1991: 462–64), which convinces many of the primacy of evidence-relative wrongdoing.

Suppose a patient has a painful but non-lethal disease. The doctor has three options: (1) give drug A, which will completely cure the patient, (2) give drug B, which will only incompletely cure the patient, and (3) give drug C, which will kill the patient. The

<sup>12</sup>Thanks to an anonymous referee for encouraging me to explicitly address this view and clarify why I reject it.

<sup>13</sup>Parfit (2011: 150).

<sup>14</sup>Parfit (2011: 151).

<sup>15</sup>For the view that fact-relative wrongdoing is primary, see Graham (2010) and Thomson (1986: 173–91). For the opposing view, see, among many others, Jackson (1991) and Parfit (2011: 150–64).

doctor knows that drug B will only incompletely cure the patient, but they do not know which of A and C will cure or kill them. Most of us hold that the doctor should choose (2) and that it would be wrong of them to choose (1) or (3) instead, even though (1) will completely cure the patient and (2) will not. But then wrongdoing must depend, at least primarily, on the available evidence, not the facts.

Even if Jackson (and others) are mistaken about this example, it still seems that fact-relative wrongdoing is not primary in distributive ethics. While it is often justified in forced choices to disadvantage evidence-relative wrongdoers relative to their victims (as in *Evil Trolley*), this is not so with respect to fact-relative wrongdoers and their victims. Few would hold, for instance, that because Jackson's doctor does something fact-relatively wrong in giving drug B, they are liable to pay compensation. Or consider another example. Didi and Dario are crossing a footbridge which their evidence suggests to be safe. Didi, who is walking in front, puts her foot on a weak spot in the bridge, causing it to collapse. It is fact-relatively wrong of Didi to put her foot on this spot. But this does not seem to weaken Didi's claim to aid relative to Dario's. If rescuers can prevent only one of them from plunging down (but not both), they should give each an equal chance of rescue.

Similar ideas are espoused in the context of defensive harming. When someone poses a threat of harm to another person as a result of fact-relatively wrong but evidence-relatively permissible behaviour, many believe that this person is not liable to defensive harm.<sup>16</sup> Suppose Ebru will cause Eli to fly into a murderous rage by sitting next to him on the bus, but Ebru has no way of knowing or foreseeing this. The fact that Ebru acts fact-relatively wrongly by sitting next to Eli on the bus does not seem to weaken her claim against harm relative to Eli's other victims' claims.<sup>17</sup> (Depending on the number of victims, there might be a lesser-evil justification for harming Ebru, and the victims themselves might have agent-relative prerogatives for defensively harming her. But neither justification invokes Ebru's fact-relative wrongdoing as the salient feature.)

There is thus good reason to think that fact-relative wrongdoing does not justify unequal treatment in forced choices. But then Beth's fact-relative wrongdoing cannot be what justifies unequal treatment in *Heroic Trolley*. We should look elsewhere for our answer. (Hereafter, deontic labels such as "(im)permissible" and "supererogatory" are used only in their evidence-relative sense.)

### 3.2. Option luck

Many philosophers believe that neutralising the effects of bad brute luck has moral priority over neutralising the effects of bad option luck.<sup>18</sup> Option luck is "a matter of how

<sup>16</sup>See, for instance, Ferzan (2012). Some disagree with the general claim. For instance, McMahan (2009: 165–66) argues that fact-relative wrongdoing can give rise to liability when the agent could reasonably foresee it as a possible outcome of their actions, even if these actions were evidence-relatively permissible. I do not explore this debate here. I merely want to suggest that the same intuition that underpins judgements about the doctor's non-liability to compensate and Didi's equal claim to aid seems to also underpin an important view about liability to defensive harm.

<sup>17</sup>This intuition is even clearer when the evidence does not merely permit the act, but requires it or makes it supererogatory. For brief discussion of the former type of case, see McMahan (2009: 165), Otsuka (2009: 382–83), and Tadros (2011: 232–35). I am not aware of existing discussions of the latter type of case, but this paper hopes to be the starting point.

<sup>18</sup>See, for instance, Arneson (2004) and Dworkin (2000).

deliberate and calculated gambles turn out – whether someone gains or loses through accepting an isolated risk he or she should have anticipated and might have declined”, while brute luck is “a matter of how risks fall out that are not in that sense deliberate gambles” (Dworkin 2000: 73).<sup>19</sup> One might think that Claire may disadvantage Beth relative to Anne in *Heroic Trolley* because Beth’s predicament is bad option luck, whereas Anne’s predicament is bad brute luck. Anne suffers bad brute luck because her predicament is not a foreseeable result of a choice of hers. Beth suffers bad option luck because it foreseeably results from her choice to try to rescue Anne. Since the attempt to rescue Anne is supererogatory, Beth could have declined the risk.

In line with this section’s general strategy, I grant that an asymmetry in the kind of luck people suffer sometimes settles who should be prioritised in a forced choice. But the asymmetry does not plausibly settle this in *Heroic Trolley*. When someone’s risk-taking is supererogatory and done for the sake of another person, the optionality of the risk-taking does not seem to justify disfavouring the agent in a forced choice.<sup>20</sup>

Suppose a terrorist is stabbing people on the street. Faisal decides to leave his hiding spot to incapacitate the terrorist, knowing he will likely incur serious injuries in the process. Faisal incapacitates the terrorist, but incurs serious injuries in doing so. When emergency medical technicians (EMTs) arrive, they find many people injured to the same degree as Faisal. They must choose whom to treat first. If the optionality of risk-taking justifies disadvantaging agents in forced choices even when their risk-taking is supererogatory and altruistic, the EMTs would have luck-based reasons for treating the brute luck victims of the attack before treating Faisal. But it seems the EMTs have no such reasons. The fact that Faisal chose to risk his injuries, and the other victims did not, does not seem to give the EMTs justification for deprioritising Faisal relative to other victims.

The idea that morality does not treat supererogatory altruistic gambles the same way it treats merely permissible self-interested gambles gets further support from a comparison with morally required gambles. Imagine a version of *Heroic Trolley* in which Beth’s chance of saving Anne is 0.99, so Beth has a duty to attempt to rescue Anne. If the 0.01 risk materialises and results in a forced choice, it seems clear that Claire may not disfavour Beth relative to Anne. Beth only does what morality asks her to do, so surely, she should get at least an equal chance of rescue.<sup>21</sup> But if Beth is off the hook when she does what is required of her, it would be strange if she is on the hook when she does *better* than is required of her. In the supererogatory case, the risky act is likewise favoured by our moral principles, and those principles likewise push Beth to perform the act. So, if option luck justifies disadvantaging someone in a forced choice, it seems to do so only when morality does not favour the optional act and does not encourage the person to perform it.

### 3.3. Autonomous preferences

A third view holds that unequal treatment of Anne and Beth in *Heroic Trolley* is justified by an asymmetry in Anne’s and Beth’s autonomous preferences. Beth makes a

<sup>19</sup>I stick to this classical formulation of the distinction. For different accounts, see, for instance, Lippert-Rasmussen (2001), Otsuka (2002), and Vallentyne (2002).

<sup>20</sup>See Eyal (2007: 3–4, 11–12) and Temkin (2003: 144–45) for similar observations. The more general idea being expressed here – that when we act for others, we are not morally on the hook should things go awry and we end up endangering those others – has recently been developed and defended by Viehoff (2022).

<sup>21</sup>See Otsuka (2009: 383) for a similar claim.



deliberate choice to risk her own life for the sake of saving Anne, so she is prepared to make a potentially huge sacrifice to attain this goal. This tells us that Beth strongly values the end of saving Anne. By contrast, nothing Anne has done indicates that Anne strongly values the end of saving Beth. Given this, one could argue that Claire should disadvantage Beth relative to Anne because she would disrespect Beth as an autonomous moral agent if she did otherwise.

The idea that it is disrespectful to interfere with the autonomous decisions of others is widely accepted in non-consequentialist moral theory. Moreover, it looks plausible in other contexts that we should distribute unavoidable harm in ways that avoid frustrating strongly valued ends, even if those distributions are suboptimal in other respects. For instance, McMahan (2005: 396) gives the example of a mother who is about to be wrongfully killed by her son but asks the potential defender not to intervene, in order to prevent her son from being harmed. It would plausibly wrong the mother to kill the son.

There are important differences between McMahan's example and *Heroic Trolley* that explain why autonomy-based reasons do not justify deprioritising Beth relative to Anne, even if they justify not saving the mother. Unlike the mother and the son, Anne and Beth are strangers to one another. This should make us question how strongly Beth really values the goal of saving Anne, since it cannot be part of Beth's long-standing projects (unless she is pursuing an altruistic project, but assume that this is not the case). Furthermore, the mother asks the potential defender not to save her, but Beth is not asking Claire not to be saved. In light of these considerations, it seems that Claire lacks sufficient evidence for concluding that Beth values the goal of saving Anne so strongly that she is willing to bear even greater risks of death for the sake of it.

One might grant this point, but insist that we can infer from the mere fact that Beth is willing to risk her life for Anne's sake that she would prefer Anne to be saved rather than herself should a forced choice arise. It might seem inconsistent for Beth to risk dying to prevent Anne from dying, while preferring Anne's death over her own in the forced choice. However, Beth's decision to risk her life to rescue Anne does not rationally commit her to prefer that Anne be put before herself in a forced choice. In attempting the rescue, Beth takes at most a 0.15 risk of dying. She takes this risk in order to increase Anne's prospect of survival by 0.4. Given these numbers, Beth can value her own life much more than Anne's and still rationally attempt the rescue.

### 3.4. Beneficiary liability

The views discussed so far all favoured saving Anne. The beneficiary liability view, if successful, favours saving Beth. The idea of beneficiary liability is advocated by several just war theorists.<sup>22</sup> These theorists argue that it is permissible to transfer some of the risk involved in a justified humanitarian intervention from intervening combatants onto imperilled civilians, because these risks are offset by the benefit that the intervention is expected to produce for the civilians on the whole. As long as the intervention still sufficiently increases the civilians' overall chance of survival, they lack a complaint against being made to share in the risks. In McMahan's words, "if [imperilled civilians] would still derive a net benefit from the defensive action even after suffering the harms that are

<sup>22</sup>See, for instance, Draper (2015: 160–65), McMahan (2010a, 2010b), and Øverland (2011). For criticism, see Christie (2018) and Eggert (2018).



an unavoidable concomitant of a successful defense, they have no valid grounds for complaint that other innocent people were not made to take a share of those harms” (2010b: 67). Other civilians, who are not imperilled and do not stand to benefit from the intervention, would have valid grounds for complaint and cannot permissibly be subjected to a share of the risk.

One might argue that Claire may subject Anne to a share of the risk that would otherwise fall on Beth in the forced choice, because Anne was expected to benefit from the risky intervention that Beth undertook for Anne’s sake. However, this argument fails. It is true that, according to the beneficiary liability view, *Beth* could have justified transferring some of the risk of her intervention onto Anne, provided that the intervention would still increase Anne’s overall chance of survival enough to be strongly in Anne’s interest. But we are evaluating whether a beneficiary liability justification is available at a time when Beth has already failed to save Anne and Anne is no longer expected to benefit from Beth’s intervention – namely, the time at which Claire is deciding whom to save in the forced choice. If Anne is no longer expected to benefit from Beth’s intervention at *this* time, then Claire cannot justify shifting risks onto Anne on the basis of Anne’s status as Beth’s expected beneficiary. Put simply, the problem is that Anne no longer fulfils the putative grounds of beneficiary liability in the relevant decision moment.<sup>23</sup>

Still, I think the beneficiary liability view is right about the following: Claire may favour Beth in the forced choice and this has something to do with the fact that Beth exposed herself to significant risk in order to improve Anne’s chance of survival. But what exactly does it have to do with this fact? How could Beth’s trying to save Anne at an earlier point in time make it permissible for Claire to shift burdens onto Anne now, when Anne is no longer an expected beneficiary of Beth’s attempt? In the remaining sections, I develop my answer to these questions. I argue that Beth’s risky attempt to save Anne places Anne under a reciprocal duty to bear risk for Beth’s sake in the forced choice, and Claire may (or even should) fulfil this duty for Anne, by favouring Beth. According to this view, it is thus a deontic fact that gives Claire permission to impose the extra risk on Anne (Anne’s presently being required to bear this risk for Beth’s sake), rather than simply the fact that, at an earlier point in time, Anne was expected to benefit from Beth’s risky rescue attempt.

#### 4. Reciprocity

I start by saying something about when we normally incur duties to reciprocate (attempts at) aiding and what is ordinarily required of us when we have such duties. Philosophers frequently appeal to reciprocity in the course of moral and political debates. For example, A. D. M. Walker (1988: 205) argues that duties of reciprocity form the basis of political obligation, David Rodin (2014) maintains that reciprocity relations ground interpersonal rights, and John Rawls suggests that his difference principle is rooted in “a conception of reciprocity” (1971: 102). However, the most detailed discussion of reciprocal duties occurs within the literature on gratitude.

<sup>23</sup>Christie (2018: 198–99) observes something similar in humanitarian interventions. If the risks shifted onto beneficiary civilians materialise and those civilians are killed or otherwise harmed by intervening combatants, we cannot justify these harms on the basis of the civilians’ beneficiary status, since the civilians no longer have this status. If Christie is right about this, it might be worth exploring in future work whether reciprocal duties could be the ground of beneficiary liability not just in *Heroic Trolley*, but also in humanitarian interventions, since an account of beneficiary liability based on reciprocal duties would avoid this problem.

Gratitude is typically understood as the called-for response to a morally outstanding (or, equivalently, exceptional) act of (attempted) benefitting by a benefactor.<sup>24</sup> When we are grateful to someone for benefitting us (or attempting to benefit us), this involves a complex disposition with affective, cognitive, and conative components.<sup>25</sup> A typical grateful person has positive feelings and beliefs about the benefactor, such as a feeling of goodwill or the belief that the benefactor is generous. Importantly, a grateful person also typically has a desire to repay the benefactor: they consider themselves “indebted” to them and are motivated to undertake reciprocating behaviours to settle this “debt”, such as returning a favour.

Many philosophers of gratitude claim that this desire to repay tracks genuine moral duties.<sup>26</sup> Like most duties, these duties are defeasible rather than absolute. For instance, when the benefactor does not want the beneficiary to reciprocate the benefitting, this plausibly defeats the beneficiary’s duty to do so. The duties are similarly defeated when the benefactor has died, or when returning a benefit would cause disproportionate harm to bystanders.

The idea that there are genuine (defeasible) duties to reciprocate accords with common sense. In our everyday moral lives, we often treat failures to reciprocate beneficial acts as wrongs. Suppose Gary routinely accepts favours from his neighbours but never does them any favours in return. Over time, Gary’s neighbours will likely come to resent Gary for his failure to reciprocate. They will come to harbour negative feelings towards him and perhaps talk badly about him behind his back. Moreover, although we typically may not use serious levels of force or coercion to enforce beneficiaries’ reciprocal duties, it seems fine to impose minor sanctions on them in a bid to get them to act. For example, it seems fine for Gary’s neighbours to impose, or threaten to impose, minor social sanctions, such as exclusion from certain social activities.

Philosophers of gratitude disagree about the precise relation between gratitude and reciprocity. It is clear that the two are not identical. We can be obligated to return a benefit even in the absence of grateful thoughts and feelings, for instance when the benefactor is someone whom we dislike. However, some think that gratitude entails reciprocity. Tony Manela (2016), for example, claims that someone who lacks a desire to reciprocate does not count as grateful. Others maintain that reciprocity is a component of justice (as opposed to gratitude) and that gratitude and reciprocity merely co-occur in response to the same acts.<sup>27</sup> It does not matter for my purposes which of these views is correct. It is enough to know that reciprocal duties are incurred in response to what I call *morally outstanding beneficial acts*.<sup>28</sup>

<sup>24</sup>Berger (1975), Manela (2015, 2019), and McConnell (1993: ch. 1, 2017). Some philosophers describe acts that call for gratitude as “morally significant” instead. Since many morally significant acts (e.g. paying one’s employees a minimal wage) do not call for gratitude, I prefer “morally outstanding” or “morally exceptional”.

<sup>25</sup>Manela (2019).

<sup>26</sup>See, for instance, Berger (1975), Camenisch (1981), Manela (2015), and Simmons (1979: 163–83). Some, such as Card (1988) and Wellman (1999), disagree.

<sup>27</sup>Fitzgerald (1998) and Weiss (1985).

<sup>28</sup>I do not take a stance on whether we can also incur reciprocal duties in response to acts that are not morally outstanding, because the argument of this paper does not depend on it. I argue below that Beth’s attempt to rescue Anne is morally outstanding, so it is sufficient for my purposes that, absent defeaters, acts of this type give rise to reciprocal duties.

#### 4.1. Morally outstanding beneficial acts

What makes a beneficial act morally outstanding? I argue that an act's exceptionality is modulated by three factors: (i) the magnitude of benefit, (ii) the quality of the benefactor's attitudes in conferring the benefit, and (iii) the costliness of the benefitting for the benefactor, where costs include effort and harm as well as risks of harm.<sup>29</sup> Each of the three factors, if realised, contributes to making a beneficial act morally outstanding.<sup>30</sup> However, the factors are not individually necessary to that end. If, in a given case, one of the factors is absent, but the other two are realised to a high degree, then the act can still be morally outstanding.

The cost-factor, benefit-factor, and attitude-factor are also not jointly sufficient for a beneficial act's being morally outstanding. They form a set of sufficient conditions only in conjunction with three necessary conditions. First, the benefactor must reasonably believe that the beneficiary welcomes (or would welcome) the beneficial act.<sup>31</sup> The actions of a stalker who keeps giving someone unwanted gifts are not morally outstanding. Second, the act must be intended to benefit. Donating a large sum to charity is not morally outstanding if the donation results from inserting the wrong account number. Third, the act must be permissible. For example, a beneficial act that causes disproportionate harm to bystanders lacks moral exceptionality even if the benefactor incurs a substantial cost and acts out of concern for the beneficiary.

Although most philosophers of gratitude believe that merely attempted benefitting can be morally outstanding even in the absence of a benefit, some dispute this.<sup>32</sup> For example, in his discussion of what makes gratitude fitting, David Carr claims that "just as we could hardly be intelligibly said to fear what made us happy, or be proud of that which made us feel ashamed, it is hard to see how we might feel genuine gratitude for what does not benefit us" (2013: 20). In a more recent paper, Carr argues that gratitude "is an appropriate response only to a good or benefit that has actually been bestowed or granted", and that gratitude is therefore inappropriate in the absence of a benefit (2021: 153).

The idea that lack of success in benefitting undermines an act's moral exceptionality does not fit with everyday language. We routinely say things like "I still owe you one" or "thanks anyway" in response to unsuccessful attempts to benefit. The idea also conflicts with common-sense morality. Suppose you need a kidney transplant. You match with a stranger volunteering through the national donor scheme, and doctors use this stranger's kidney for your transplant. Unfortunately, the transplant is unsuccessful. Still, it seems you ought to be grateful to the stranger for *trying* to benefit you at significant cost to themselves, and motivated to return a benefit.

<sup>29</sup>The assumption that risks, even when unmaterialised, count as costs in themselves is controversial but sufficiently widely accepted for my purposes here.

I take (i) and (iii) to be objective in the sense that they refer to benefits and costs that in fact materialise. The magnitude of the benefit that the agent *reasonably believes* they are bestowing on the beneficiary and the magnitude of the cost that the agent *reasonably believes* they are taking on do still bear on the act's moral exceptionality. However, they do so not by improving on (i) or (iii), but by improving on (ii).

<sup>30</sup>Several philosophers of gratitude, such as McConnell (1993: ch. 1) and Simmons (1979: 170–79), identify similar factors as the bases of gratitude.

<sup>31</sup>This requirement is weaker than the possible alternative requirement that the beneficiary receives the benefit voluntarily. Even if the beneficiary has not consented to being benefitted, or invited or requested the benefitting, a beneficial act can be morally outstanding and give rise to a duty of reciprocity.

<sup>32</sup>For the view that gratitude can be owed even to would-be benefactors, see Berger (1975: 299) and McConnell (1993: 42–44).

One might object that even if someone fails to confer the benefit they intended to confer, the attempt to benefit is itself beneficial. It shows that the person cares more for us than we had previously apprehended, and that makes us feel good. However, merely attempted benefitting is not always beneficial for this reason. Reconsider the kidney transplant example. The stranger whose kidney was unsuccessfully transplanted did not have special care for you in particular; they wanted to help *whomever* they would match with through the donor scheme. Personal affection also cannot explain why gratitude can be due in response to merely attempted benefitting by someone whom you already know cares a lot for you, such as a parent or a spouse. If such a person fails to confer the intended benefit on you, you do not learn anything new about their affection for you. Still, it seems you can owe them a benefit in return.

When we have a duty to reciprocate merely attempted benefitting, it is typically because the act is morally outstanding. As indicated above, beneficial acts can be morally outstanding even if one of the factors contributing to moral exceptionality is absent. In the case of merely attempted benefitting, the absence of a benefit can be offset by a substantial sacrifice and excellent attitudes from the agent.

#### 4.2. Reciprocating aid

So far, I have suggested that reciprocal duties are incurred in response to morally outstanding beneficial acts, and I have given an account of what makes beneficial acts morally outstanding. Before returning to *Heroic Trolley*, we also need to know what we ought to do when we are under a duty to reciprocate a morally outstanding act of (attempted) benefitting. It is difficult to say anything specific about this. Much depends on the particular context in which the duty arises: a return that is fitting in one situation might not be in another. Still, I will make some general observations about what is required by one subset of reciprocal duties, namely those incurred in response to morally outstanding acts of (attempted) aiding by people whom we do not know (or do not know well). It is important to keep in mind that the content of these duties might, and likely does, differ from the content of our duties to reciprocate acts of aiding by *intimates*, or acts of *treating* (understood, loosely, as acts that do not fulfil a need).

When we ought to reciprocate a stranger's morally outstanding act of (attempted) aiding, we should normally benefit *the stranger*, rather than someone else. For example, if a stranger saves your life, saving a third party's life does not discharge your reciprocal duty to the stranger. It is *the stranger* to whom you are indebted, in virtue of what they did for you, and settling this debt requires paying back the benefit rather than paying it forward. There are, of course, exceptions. For instance, if you know that the stranger prefers that you reciprocate their act of saving by helping third parties, then that is what you ought to do. But absent such special conditions, discharging reciprocal duties requires benefitting the benefactor in particular.

If the initial act of (attempted) aiding is not extremely costly for the benefactor, we should normally reciprocate by performing an act of (attempted) aiding that is roughly as outstanding as the initial act. Returning aid that is much more or much less beneficial or costly than the initial act is typically inappropriate.<sup>33</sup> For example, if a stranger helps you carry luggage and you offer your plane ticket in return, the return would overstate

<sup>33</sup>One exception might be situations in which the beneficiary is significantly less well-off than the benefactor and unable to make an equal sacrifice. In those situations, it is plausibly an equal *marginal* sacrifice that is demanded by reciprocity, as Becker (1990) also suggests.

the significance of what the stranger did for you. Conversely, if your neighbour spends an entire weekend helping you move into your new home, it would be improper to merely say “thank you” and consider your debt settled. This would seem to trivialise the value of what the neighbour did for you. In addition to thanking the neighbour, it seems you should seek to return aid of a kind that approximates the significance of the aid you received.

What about cases in which the initial act of (attempted) aiding is very costly to the benefactor? Are beneficiaries required to take on costs of comparable magnitude even then? I do not think so. Suppose you are drowning and a stranger saves you, incurring a cost that significantly exceeds what morality asks us to bear to save someone – for example, a shark bites off both their legs. The next month, the rescuer is about to be hit by a truck. You can push them out of harm’s way at the cost of the truck’s crushing both your legs. Although this cost is no greater than the cost that the rescuer incurred for your sake, it does not seem that you are under a duty to save your rescuer at this cost. There seems to be an upper limit to the magnitude of cost that we can be required to reciprocally incur for the sake of our benefactors.

However, this does not mean that we cannot be required to reciprocally incur any cost that would ordinarily be supererogatory. Suppose you can prevent your rescuer from being hit by the truck in a way that involves taking on only a *risk* of the truck’s crushing your legs, and this risk just exceeds what morality would ordinarily ask us to bear to save others. I believe that you would be required to save your rescuer at this risk, even though other people, who are not indebted to your rescuer, would not be. So even if there is an upper limit in place, reciprocal duties increase the magnitude of cost that we can be required to incur to save others from harm: if we can ordinarily be required to save someone from some harm at cost  $c$ , we can be required to reciprocally save our rescuer from that same harm at some cost greater than  $c$ .

One might resist this claim by pointing out that, intuitively, the rescuer may not *demand* that you reciprocally save them at this normally supererogatory risk. And if they may not demand this, then you are not actually required to save them at this risk.

Two replies. First, I disagree that the rescuer may not demand that you reciprocally save them in this scenario. Suppose the rescuer realises that you intend not to rescue them. If they would respond to this by saying “Save me! You owe me!”, that would not seem inappropriate or impermissible to me. Second, even if it were true that your rescuer may not demand that you save them in our scenario, this does not show that you are not required to save them. According to most philosophers of gratitude, benefactors may never demand reciprocal benefits from their beneficiaries, not even when the beneficiaries uncontroversially owe them those benefits.<sup>34</sup> So we cannot infer from the fact that the rescuer may not demand that you save them at a normally supererogatory risk that you are not under a duty to save them at this risk.

## 5. Reciprocity in *Heroic Trolley*

I now argue that Claire’s favouring Beth is justified by an asymmetry in the duty to reciprocate. The argument builds on three claims: (1) Beth’s attempt to rescue Anne is morally outstanding and places Anne under a reciprocal duty, (2) Anne ought to reciprocate Beth’s attempt by bearing risks for Beth’s sake in the forced choice, and (3) because Anne is unable to fulfil this duty herself, and Claire will either fulfil

<sup>34</sup>See, for instance, Berger (1975), Camenisch (1981), and Manela (2015).

Anne's duty or ensure that the duty goes unfulfilled, Claire is permitted to fulfil it by imposing the required risks on Anne for Beth's sake. After defending these three claims, I also suggest that (4) because the stakes of Anne's duty being fulfilled are high, Claire is probably *required* to impose the required risks on Anne for Beth's sake.

Let us start by establishing that Beth's attempt to save Anne is morally outstanding. In Sections 2 and 3.1, I argued that the attempt is permissible and intended to benefit Anne, and that it is reasonable for Beth to believe that Anne would welcome it. This means that the necessary conditions for the attempt's being morally outstanding are fulfilled. The attempt also realises the three factors that contribute to an act's being morally outstanding. Beth attempts to benefit Anne willingly and out of concern for Anne, and she takes on a significant cost for Anne's sake by exposing herself to a non-trivial risk of death. In addition, the benefit Beth attempts to confer on Anne – saving Anne's life – is large. Beth does not succeed in conferring this benefit, but I suggested in the previous subsection that merely attempted benefitting can be morally outstanding if the absence of a benefit is compensated by a large sacrifice and excellent attitudes from the benefactor. Given the goodness of Beth's attitudes and the magnitude of her sacrifice, such compensation is plausibly achieved here.

If Beth's attempt to save Anne is morally outstanding, then unless there are defeating conditions in place, Anne incurs a reciprocal duty in response to it. No defeating conditions seem to be in place. Beth is still alive when Claire is deciding how to distribute the chances, Beth has not expressed that she does not want Anne to reciprocate the attempt, and there is no reason to think that Anne will harm any bystanders in returning a benefit. Moreover, since Beth has not communicated that she wants Anne to pay the benefit forward, rather than back, Anne ought to return a benefit to Beth in particular.

It is true, however, that Beth's attempt is of a kind that we are rarely able to fittingly return. It is not probable that someone who improves our chance of survival becomes similarly imperilled later. It is even less probable that if this person becomes similarly imperilled later, we are in a position to improve their chance of survival. But unusually, in *Heroic Trolley*, the person who improved Anne's chance of survival, Beth, does become similarly imperilled. Of course, Anne is unable to reciprocally improve Beth's chance of survival, since she cannot act or communicate in the forced choice. Still, if Anne *were* in a position to improve Beth's chance of survival, she would be required to do so. If Anne were in the decision-making position instead of Claire, or if she could authorise Claire to act on her behalf, she would be required to arrange that Beth bears less than half of the risk in the forced choice, by taking on more than half herself.

In *Heroic Trolley*, Anne is incapacitated and cannot reciprocally improve Beth's chance of survival. But Claire can. Claire can see to it that Anne bears the extra risks for Beth's sake, thereby fulfilling Anne's duty to Beth.<sup>35</sup> This would not mean that Claire is acting *for* Anne, or on Anne's behalf. My claim is that Claire can fulfil Anne's duty to Beth even if she acts merely on her own behalf. This is because, unlike duties of gratitude, duties of reciprocity typically lack an attitudinal component. While duties of gratitude seem to require that the beneficiary has certain feelings towards the

<sup>35</sup>Strictly speaking, if Anne is incapacitated, she is not required to do anything. But to get my argument going I need it to be true only that the grounds of the duty to bear risks for Beth's sake are satisfied, such that if Anne were not incapacitated, Anne would be under the duty. For brevity, I just say that Anne is under the duty.

benefactors and expresses those in their actions, reciprocal duties seem to require only that a certain outcome is brought about – namely, that the benefactor is proportionally benefitted at the beneficiary’s expense. And while it is plausible to think that third parties typically cannot fulfil the attitudinal components of other people’s duties, there is no similar obstacle to fulfilling the outcome-oriented components of other people’s duties.

A comparison with duties to apologise and duties to compensate might be helpful. Like duties of gratitude, duties to apologise seem to have an attitudinal component. We typically believe that wrongdoers should feel remorse about their wrongdoing and that their apologies should express this feeling. Plausibly, only the duty-bearers themselves can fulfil these requirements. But duties to compensate, like duties to reciprocate, seem to require only that a certain outcome is brought about: that the victim is proportionally benefitted at the wrongdoer’s expense. And there is no similar obstacle to third parties’ fulfilling this requirement.

Some might still deny that Claire can fulfil Anne’s reciprocal duty to Beth. They might argue, for instance, that the duty requires not that Anne *bears* risks for Beth’s sake in the forced choice, but, more specifically, that Anne *takes on* those risks for Beth’s sake. On such a view, Anne’s agency is built into the content of the duty, so that only she can fulfil it.

The view is unpersuasive. Suppose Claire makes Anne bear the required risks for Beth’s sake. It seems Beth could not expect any more benefits from Anne, and Anne would not be required to make any further sacrifices for Beth’s sake. This means that Anne is no longer required to reciprocate Beth’s attempt to save her. Her duty to Beth has been fulfilled.

Even if one accepts that Claire *can* fulfil Anne’s duty to Beth, one might still hold that this would have *no value*. Or, put differently, that the value of the duty’s fulfilment derives entirely from the value of Anne’s being the agent. But this is not plausible. There is at least one other value that it also realises: Beth’s getting the chances of survival that she is owed. And for this value to be realised, it does not seem to matter *who* fulfils the duty. What matters is *that* it gets fulfilled, so that Beth gets the chances of rescue she is owed. This does not mean that there is no special value realised in Anne’s being the agent; just that it is not the only value at stake in the duty.

Suppose I am right that Claire can fulfil Anne’s duty to Beth and that this would not undermine the value of the duty’s fulfilment. Would it be *permissible* for Claire to do so? I believe so. When others are unable to fulfil their duties, and we are forced to choose between fulfilling the duties and ensuring that the duties will not be fulfilled, we may ordinarily do the former.<sup>36</sup> Suppose you know that Hana promised Hiro that he could borrow her bike today. When Hiro turns up to collect the bike, you realise

<sup>36</sup>Tadros (2011: 131–36) presents a similar argument in the context of permissions to use others as means. In his view, if an incapacitated person would be required to use their body for the sake of some end, and that duty would be enforceable, then others may use the person’s body in that way for the sake of that end. Tadros’ argument includes the premise that the duty is enforceable. Mine does not. I argue that third parties may fulfil other people’s duties to reciprocate when those people are unable to fulfil the duties themselves and third parties must choose between doing something that fulfils the duties and doing something that ensures the duties will not be fulfilled. As I explain below, this is compatible with third parties’ being prohibited from fulfilling other people’s duties to reciprocate when those people are able to do it but unmotivated or unwilling. So the scope of my argument in this paper is limited. Still, I argue in Eskens (forthcoming) that expressive duties, including duties to express gratitude by reciprocating, are in fact enforceable, so I believe that the argument’s scope can be extended even to cases in which duty-



that Hana has forgotten and that she is unreachable for the day. You have two options: fulfil Hana's duty to Hiro, by giving Hiro the bike, or ensure that the duty will not be fulfilled, by refraining from giving it to him. Since you *know*, rather than merely suspect, that Hana ought to lend her bike to Hiro, it seems you may do what is compatible with Hana's duty's being fulfilled: give Hiro the bike.

Claire is permitted to fulfil Anne's duty to Beth for the same reason you are permitted to fulfil Hana's duty to Hiro. Just as you have two options with respect to the fulfilment of Hana's duty to Hiro, so Claire has two options with respect to the fulfilment of Anne's duty to Beth. Claire can either fulfil the duty, by favouring Beth, or make it certain that it will not be fulfilled, by refraining from favouring Beth. If Claire *knows* that Anne is required to bear risks for Beth's sake in the forced choice, rather than merely suspecting it, it seems Claire may do what is compatible with Anne's duty being fulfilled: favour Beth.

Three clarifications. First, even if Claire may fulfil Anne's duty to Beth when Anne is unable to do it herself, this does not mean that there is no special value in Anne's fulfilling the duty herself if she were able to. If either Anne or Claire can fulfil Anne's duty, it is probably more valuable if Anne does it than if Claire does it. Second, if Claire may fulfil Anne's duty to Beth when Anne is unable to do it herself, this does not imply that Claire may also fulfil Anne's duty to Beth if Anne were able to do it herself but unwilling or unmotivated to do so. It could be that Claire is prohibited from interfering with Anne's violating her duty even if Claire is permitted to fulfil the duty when Anne is merely unable to fulfil it. Third, if Claire may fulfil Anne's duty to Beth without Anne's consent when Anne is unable to fulfil the duty and unable to consent, this does not mean that Claire would not be required to seek Anne's consent if she were able to. In giving Anne the chance to consent, Claire would enable Anne to take control of the decision and over what happens to her, and giving persons this sort of control over their lives is important.<sup>37</sup>

If my argument is successful, Claire may favour Beth because she may act compatibly with Anne's duty to Beth being fulfilled. How much may she favour Beth? I am not sure. A precise answer to this question requires a precise account of the magnitude of risk Anne is required to reciprocally incur for Beth's sake, since this fixes the amount of risk Claire may impose on Anne. I do not at present have such an account. Still, it seems unlikely that Claire may favour Beth so much that Anne is left with no chance of rescue at all, because it is unlikely that Anne is required by reciprocity to take on *all* the risk in the forced choice. That would seem disproportionate compared to the initial benefitting and exceed the upper limit on the cost that we can be required to reciprocally incur for our benefactors' sakes (see Section 4.2). So, my view probably permits only moderate favouring: Claire may give Beth a greater chance of rescue than Anne, but she may not simply save Beth.

If Claire's choice is not dispositive – she can distribute the chances of being saved rather than simply saving either Anne or Beth – this creates no difficulties. I am not sure what the implications are when Claire's choice *is* dispositive. Does the fact that Claire is permitted to give Beth a greater chance than Anne in a non-dispositive choice give Claire reason to save Beth rather than Anne in a dispositive choice? Or does the fact

---

bearers are able but unmotivated or unwilling to fulfil their duties of reciprocity. This is, however, not a claim I can defend in this paper.

<sup>37</sup>For the view that the value of consent in this type of case is rooted in control, see Tadros (2011: 133–34).

that she would not be permitted to simply save Beth in a non-dispositive choice prevent such reason from arising? These are important questions. I hope to explore them in future work, but I cannot answer them here.

I have argued that Claire may favour Beth by giving her a greater chance of rescue than Anne. This is sufficient to account for the intuition that the presumption of equality is overturned in *Heroic Trolley*. But my argument might support a stronger conclusion: Claire is *required*, rather than merely permitted, to favour Beth. This is because high stakes attach to the duty's fulfilment, namely Beth's getting the chances of rescue she is owed. The fact that there are these high stakes makes it difficult to see why Claire would have discretion over whether the duty gets fulfilled. We would say the same, I think, about a high-stakes version of the bike example. If Hiro relies on the bike to get to an important job interview in time or to visit his dying grandmother, it does not seem up to you to decide whether he gets to borrow it. You seem required to do what fulfils the duty, and prohibited from doing what ensures it goes unfulfilled.

Some might want to resist the stronger conclusion, for instance by claiming that it would be unfitting to blame Claire for refraining from favouring Beth. It does not really matter for my purposes whether the weaker or stronger conclusion is correct. The weaker conclusion that Claire is permitted to favour Beth is already sufficient to account for the intuition I wanted to account for: that Claire is not required to give Anne and Beth equal chances of being saved.

Still, it is worth saying something in response to the objection. It assumes a strong pre-theoretical intuition that Claire would not be blameworthy for refraining from favouring Beth. But I doubt that most readers have such an intuition, because I doubt that they have strong pre-theoretical intuitions about what exactly Claire is permitted or required to do. The only intuition I took for granted in the beginning is that Claire is not required to treat Anne and Beth equally. This left it open whether the inequality would favour Anne or Beth and whether giving equal chances would still be permissible. I suspect that readers' intuitions pull in different directions. But if we lack clear intuitions about what *exactly* Claire is permitted or required to do, then we must also lack clear intuitions about what *exactly* Claire would be blameworthy for (not) doing, since those are two sides of the same coin. Because the answer to what exactly Claire may or should do must thus come from theory rather than intuition, I believe we should accept the verdict supported by our best theory. I have argued that the reciprocity account is the best theory.

## 6. Conclusion

This paper identified a gap in the literature on distributive ethics: existing accounts of the fair distribution of benefits and burdens in forced choices cannot solve a case like *Heroic Trolley*, in which a forced choice between lives is caused by one of the imperilled persons' supererogatory attempt to rescue the initially imperilled person. I defended a reciprocity-based view that fills this gap. This view says that when the initially imperilled person has a duty to reciprocate the unsuccessful rescuer's attempt to save them, by bearing risks for their sake in the forced choice, the decision-maker may (or even should) see to it that this duty is fulfilled if the initially imperilled person cannot fulfil it themselves, thus favouring the unsuccessful rescuer.

**Acknowledgements.** I am very grateful to Gunnar Björnsson, Helen Frowe, and Jonas Haeg for detailed feedback on multiple drafts of the paper and many helpful conversations about the material in it. I also thank Alfred Archer, Andrea S. Asker, Stina Björkholm, Joseph Bowen, Julia Driver, William Gildea,

Robert J. Hartman, Lisa Hecht, Kieran Oberman, Meradjuddin Khan Oidermaa, Theron Pummer, H. Orri Steffánson, Jay Wallace, and two anonymous referees for their comments and suggestions. Versions of the paper were presented at the REAPP Graduate/Early Career Conference, the Society for Applied Philosophy Annual Conference, the Pavia Graduate Conference in Political Philosophy, the SCEWP Annual Graduate and Early Career Reading Retreat, the Swedish Congress for Philosophy, and Stockholm University.

**Financial support.** This research was supported by the Knut and Alice Wallenberg Foundation, grant no. 1521101.

## References

- Arneson, R. 2004. Luck Egalitarianism Interpreted and Defended. *Philosophical Topics* 32: 1–20.
- Becker, L. 1990. *Reciprocity*. Chicago: University of Chicago Press.
- Berger, F. 1975. Gratitude. *Ethics* 85: 298–309.
- Camenisch, P. F. 1981. Gift and Gratitude in Ethics. *Journal of Religious Ethics* 9: 1–34.
- Card, C. 1988. Gratitude and Obligation. *American Philosophical Quarterly* 25: 115–27.
- Carr, D. 2013. Varieties of Gratitude. *Journal of Value Inquiry* 47: 17–28.
- Carr, D. 2021. From (Apparently) Feeling to Being Grateful. *Journal of Value Inquiry* 55: 145–54.
- Christie, L. 2018. Distributing Death in Humanitarian Interventions. In *Who Should Die? The Ethics of Killing in War*. Edited by Bradley Jay Strawser, Ryan Jenkins and Michael Robillard. Oxford: OUP: 186–201.
- Draper, K. 2015. *War and Individual Rights: The Foundations of Just War Theory*. Oxford: OUP.
- Driver, J. 1997. The Ethics of Intervention. *Philosophy & Phenomenological Research* 57: 851–70.
- Dworkin, R. 2000. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge, MA: HUP.
- Eggert, L. 2018. Harming the Beneficiaries of Humanitarian Intervention. *Ethical Theory & Moral Practice* 21: 1035–50.
- Eskens, R. (forthcoming). Expressive Duties are Demandable and Enforceable. In *Oxford Studies in Normative Ethics*, Volume 14. Edited by Mark Timmons. Oxford: OUP.
- Eskens, R. (manuscript). Gratitude and Rights. Manuscript.
- Eyal, N. 2007. Egalitarian Justice and Innocent Choice. *Journal of Ethics & Social Philosophy* 2: 1–18.
- Ferzan, K. K. 2012. Culpable Aggression: The Basis for Moral Liability to Defensive Killing. *Ohio State Journal of Criminal Law* 9: 669–98.
- Fitzgerald, P. 1998. Gratitude and Justice. *Ethics* 109: 119–53.
- Frowe, H. 2021. The Limited Use View of the Duty to Save. In *Oxford Studies in Political Philosophy*. Volume 7. Edited by David Sobel, Peter Vallentyne and Steven Wall. Oxford: OUP: 66–99.
- Graham, P. 2010. In Defense of Objectivism about Moral Obligation. *Ethics* 121: 88–115.
- Jackson, F. 1991. Decision-Theoretic Consequentialism and the Nearest and Dearest Objection. *Ethics* 101: 461–82.
- Lippert-Rasmussen, K. 2001. Equality, Option Luck, and Responsibility. *Ethics* 111: 548–79.
- Manela, T. 2015. Obligations of Gratitude and Correlative Rights. In *Oxford Studies in Normative Ethics*, Volume 5. Edited by Mark Timmons. Oxford: OUP: 151–70.
- Manela, T. 2016. Gratitude and Appreciation. *American Philosophical Quarterly* 53: 281–94.
- Manela, T. 2019. Gratitude. *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition). Edited by Edward N. Zalta. <https://plato.stanford.edu/archives/spr2015/entries/gratitude/>.
- McConnell, T. 1993. *Gratitude*. Philadelphia: Temple University Press.
- McConnell, T. 2017. Gratitude, Rights, and Moral Standouts. *Ethical Theory & Moral Practice* 20: 279–93.
- McMahan, J. 2005. The Basis of Moral Liability to Defensive Killing. *Philosophical Issues* 15: 386–405.
- McMahan, J. 2009. *Killing in War*. Oxford: OUP.
- McMahan, J. (2010a). The Just Distribution of Harm between Combatants and Non-combatants. *Philosophy & Public Affairs* 38: 342–79.
- McMahan, J. (2010b). Humanitarian Intervention, Consent, and Proportionality. In *Ethics and Humanity: Themes from the Philosophy of Jonathan Glover*. Edited by N. Ann. Davis, Richard Keshen and Jeff McMahan. Oxford: OUP: 44–71.
- Otsuka, M. 2002. Luck, Insurance, and Equality. *Ethics* 113: 40–54.
- Otsuka, M. 2009. Moral Luck: Optional, Not Brute. *Philosophical Perspectives* 23: 373–88.

- Øverland, G.** 2011. High-Fliers: Who Should Bear the Risk of Humanitarian Intervention? In *New Wars and New Soldiers: Military Ethics in the Contemporary World*. Edited by Paolo Tripodi and Jessica Wolfendale. Farnham: Ashgate: 69–86.
- Parfit, D.** 2011. *On What Matters*, Volume 1. Oxford: OUP.
- Rawls, J.** 1971. *A Theory of Justice*. Cambridge, MA: HUP.
- Rodin, D.** 2014. The Reciprocity Theory of Rights. *Law & Philosophy* 33: 281–308.
- Simmons, A. J.** 1979. *Moral Principles and Political Obligations*. Princeton: PUP.
- Tadros, V.** 2011. *The Ends of Harm: The Moral Foundations of Criminal Law*. Oxford: OUP.
- Temkin, L.** 2003. Exploring the Roots of Egalitarian Concern. *Theoria* 69: 125–51.
- Thomson, J. J.** 1986. *Rights, Restitution, and Risk*. Edited by William Parent. Cambridge, MA: HUP.
- Vallentyne, P.** 2002. Brute Luck, Option Luck, and Equality of Initial Opportunities. *Ethics* 112: 529–57.
- Viehoff, D.** 2022. Legitimate Injustice and Acting for Others. *Philosophy & Public Affairs* 50: 301–74.
- Walker, A. D. M.** 1988. Political Obligation and the Argument from Gratitude. *Philosophy & Public Affairs* 17: 191–211.
- Weiss, R.** 1985. The Moral and Social Dimensions of Gratitude. *The Journal of Southern Philosophy* 23: 491–501.
- Wellman, C.** 1999. Gratitude as a Virtue. *Pacific Philosophical Quarterly* 80: 284–300.