

University of Groningen

The Problem of Collective Harm

Hindriks, Frank

Published in:
Erasmus Journal for Philosophy and Economics

DOI:
[10.23941/EJPE.V1711.798](https://doi.org/10.23941/EJPE.V1711.798)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2024

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Hindriks, F. (2024). The Problem of Collective Harm: A Threshold Solution. *Erasmus Journal for Philosophy and Economics*, 17(1), 209-233. <https://doi.org/10.23941/EJPE.V1711.798>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

The Problem of Collective Harm: A Threshold Solution

FRANK HINDRIKS
University of Groningen

Abstract: Many harms are collective: they are due to several individual actions that are as such harmless. At least in some cases, it seems impermissible to contribute to such harms, even if individual agents do not make a difference. The Problem of Collective Harm is the challenge of explaining why. I argue that, if the action is to be permissible, *the probability of making a difference to harm must be small enough*. This in turn means that both the probability of harm and the probability of avoiding harm have to remain below the corresponding threshold probabilities. I compare this threshold probability account to proposals that revolve around difference-making, *NESS* causation and security dependence, and I argue that they fail for reasons of scope. For instance, a moral principle that invokes *NESS* causation prohibits so many actions that compliance with it would have a stifling effect on human life.

Keywords: collective harm, making a difference, *NESS*, probability of making a difference, risk, security-dependence

JEL Classification: A130, D71

You are out on a Saturday night heading towards your favorite dance club. Recently, you have started noticing that people sometimes faint when it is very busy. So, you now have misgivings about its door policy and you wonder whether you should go somewhere else. However, you like this club much better than the others. And you tell yourself that your presence will, in all likelihood, not make a difference. If someone faints tonight, they will do so irrespective of whether you are there. This argument soothes your conscience. So, you decide to enter the club after all. Is it permissible for you to do so?

All individual actions in **CLUB**, as I call this example, are harmless on their own. Yet, collectively, they generate a harmful outcome. This is why

it exemplifies the problem of collective harm.¹ The fact that the outcome is harmful suggests that one or more people act impermissibly. But the fact that none of the actions is harmful as such counts against this. Presumably, this is why our intuitions about particular cases sometimes vacillate (Gunnemyr and Touborg 2023a). However, there is a third alternative. It could be that contributing to a collective harm is permissible up to a point. This seems plausible in the case of CLUB. Overcrowding should be prevented. Yet, it would go too far to prohibit clubbing altogether.

To account for this, I invoke the probability of making a difference to a harm. As I have argued elsewhere, this probability can be so high as to make the action impermissible (Hindriks 2022, 2023). Conversely, to be permissible, the probability that an action makes a difference to a harm must remain below a threshold. I call this ‘the threshold probability account’.²

In this paper, I compare this threshold account to proposals that revolve around three different conceptions of causation: difference-making; *NESS* causation (Wright 1988; Braham and Hees 2012); and security dependence (Gunnemyr and Touborg 2023a, 2023b). The corresponding moral principles feature a *pro tanto* obligation not to make a difference to harm, make a causal contribution to harm, and make it more secure. I criticize them for reasons of scope. The first is too permissive. The second prohibits so many actions that compliance with them would have a stifling effect on human life. And both the second and the third support obligations even when there are not enough others willing to contribute so as to successfully prevent the harm. The threshold probability account performs better in these respects.

Section I introduces the problem of collective harm along with the notion of difference-making. In sections II and III, I criticize *NESS* causation and security dependence. And I present and defend the threshold probability account in section IV.

¹ Nefsky (2011) uses this term. A more general version of it also encompasses collective benefits, which Nefsky (2019) calls this ‘the inefficacy problem.’ Elsewhere, I call it ‘the problem of insignificant hands’ (Hindriks 2022). This serves to highlight the similarities and differences with the problem of many hands (Thompson 1980).

² Kagan (2011) also invokes the probability of making a difference to solve the problem of collective harm, although his account is consequentialist. As I discuss in section 4.2, the threshold probability account is inconsistent with act consequentialism. See also Hindriks (2022).

I. DIFFERENCE-MAKING

A collective harm is a harmful outcome that is due to several individual actions, each harmless on its own. The problem of collective harm concerns the question whether and when it is impermissible to contribute to such a harm. An initially plausible answer is that there is a *pro tanto* obligation not to do so for reasons of harm exactly if the action causes it. Furthermore, causation can plausibly be understood in terms of making a difference. At this point, the question arises as to what it means for an action to make a difference to an outcome.

The difference-making conception of causation is commonly explicated in terms of the following two counterfactuals (Lewis 1973; List and Spiekermann 2013):³

C is a cause of *E* exactly if:

- (i) If *C* occurred (*ceteris paribus*), then *E* would occur.
- (ii) If *C* did not occur (*ceteris paribus*), then *E* would not occur.

The underlying idea is that effects track their causes. The first condition represents positive tracking, the second one negative tracking. The *ceteris paribus* clauses reveal that a cause is not, as such, necessary or sufficient for an effect. It must occur in certain circumstances for this to be the case.

However, an action need not be sufficient for an outcome in order to make a difference to it. Outcomes can be collective and depend on several actions. In such cases, the action cannot be sufficient as such. But it can be jointly sufficient for the outcome together with the other contributions. It follows that, for an action to make a difference to an outcome, it must be individually necessary and jointly sufficient for it in the relevant circumstances. At this point, it is important to notice that an action can make a difference to a harm by bringing it about or by preventing it. In light of this, I propose to formulate what I call ‘the Difference-making Principle’ as follows (*DP*):

(*DP*) There is a *pro tanto* obligation not to bring about harm.

By bringing about harm, I mean that the harm materializes due to an action that is individually necessary and jointly sufficient for it. Such an

³ List and Spiekermann add that, in condition (i), ‘E is present in all closest possible worlds (relative to the actual world) in which C is present’, while in condition (ii), ‘E is absent in all closest possible worlds (relative to the actual world) in which C is absent’ (2013, 636).

action makes a difference to harm. For an agent to have an all-things-considered obligation not to do so, she should not have an overriding obligation.

DP can be illustrated in terms of a version of *CLUB*, the example mentioned in the introduction. Your favorite club is already very busy. You are the only one left who wants to enter. If you do, the club becomes overcrowded such that one or more people will faint. In this situation, your action makes a difference. This means that, in the absence of a special justification, it is impermissible for you to enter for reasons of harm. Compare this to the situation in which there is one person less in front of you or one person behind you. In both these situations, what you do does not make a difference. Either the harm would not have occurred, or it would have been overdetermined. Hence, it is permissible for you to enter. Or if it is not, this is due to considerations other than harm.

As this example illustrates, the presence or absence of one individual agent changes the verdict that *DP* delivers. This is not very plausible. In particular, the difference-making account of causation is often criticized for giving the wrong answer in cases of overdetermination (Wright 1985; Kagan 2011; Braham and Hees 2012; Nefsky 2017). A famous example involves two assassins who shoot a politician. Each of their shots would have been fatal. Neither makes a difference. Yet, what each of them did was wrong (*ASSASSIN*).⁴ Consider, also, a species of fish that has gone extinct due to overfishing. Presumably, it was wrong for someone to eat this kind of fish even if they did not make a difference (*FISH*). These examples reveal that the difference-making account is too strong: it identifies too few events as causes. Hence, *DP* is too permissive.

According to the difference-making account, causation is a matter of counterfactual dependence. In the next two sections, I go on to consider weaker conceptions of counterfactual dependence, to wit *NESS* causation and security. They support more restrictive moral principles.

⁴ Cases of pre-emption also pose a problem for difference-making. Suppose that one assassin kills a politician. Had she been unsuccessful, another assassin would have killed him. Because of this second assassin, the first one does not make a difference to the politician's death.

II. NESS CAUSATION

II.1. THE NESS PRINCIPLE

According to the difference-making conception, a cause is (individually) necessary and (jointly) sufficient for its effect. A weaker conception of counterfactual dependence considers sets of actions that are minimally sufficient for the occurrence of an event. The claim is that an action causes an event, or makes a causal contribution to it, exactly if it is an element of such a set (Wright 1985; Braham and Hees 2009). In other words, an action makes a causal contribution to an outcome exactly if it is a Necessary Element of a Set of actions that is Sufficient for it (*NESS*). That action is a *NESS* condition for the outcome (Braham and Hees 2012, 615).

NESS causation now offers a classical solution for the problem of over-determination. Consider ASSASSIN. Although neither of the shots of the two assassins is necessary for the politician's death, each of them is a *NESS* condition for it (Wright 1985; Braham and Hees 2009). Hence, both made a causal contribution to the death of the politician. This account can be used to reformulate the principle against harm. The thing to do is replace difference-making in *DP* by *NESS* causation. This results in what I call the **NESS PRINCIPLE (NP)**:⁵

NP There is a *pro tanto* obligation not to perform an action that can be a *NESS* condition for harm.

Below I argue that this principle is too simple. Yet, it is important to consider *NP* as it captures the intuition, which some people have, that contributing to a collective harm is universally prohibited.

Walter Sinnott-Armstrong (2005) considers someone who wants to go for a drive in a gas guzzling SUV on a Sunday afternoon just for fun. By doing so, they will emit greenhouse gases, for no good reason (*JOYGUZZLING*). In this situation, it is exceedingly likely that the agent does not make a difference to climate harm. If so, the principle against harm does not prohibit the action. Yet, one might think that joyguzzling is impermissible. What is more, there might be a universal prohibition against it. By this, I mean that it is prohibited to do so, always and for everyone,

⁵ It has been argued that, to support plausible verdicts in forward-looking cases, the *NESS* condition must be combined with an account of the laws that apply or the causal mechanisms that are operative (Halpern 2008; Wright 2011). I proceed here without committing myself to a particular account, as it is unlikely to affect my argument.

except if there is a special justification. As *NP* applies to all actions that can contribute to harm, it prohibits joyguzzling. So, it can accommodate this idea.

However, as far as I know, nobody has defended this account. And for good reason, as I go on to argue.

II.II. THE STIFLING OBJECTION

The problem is that, as the future is open-ended, there are many more possible harms to reckon with than will materialize. This means that, if all contributions to all possible harms are prohibited, an awful lot of actions will be impermissible. In particular if people were to comply with the principle against harm interpreted in this strict a manner, it would have a stifling effect on human life. And it is implausible that morality has such a debilitating effect. I call this the **STIFLING OBJECTION**.

By way of illustration, consider the following three examples. Someone who enters a club can thereby make a contribution to someone's fainting. Cutting down a tree on a mountain can be a *NESS* condition for a lethal mudslide. And by eating a fish, you may make a contribution to the extinction of the species. Suppose that *NP* is valid. Then, it is impermissible to engage in these activities, at least in the absence of special justifications. Now, overcrowding should be prevented. Yet, it would go too far to prohibit clubbing altogether. This would mean that all clubs will have to shut down. The same holds for fishing and logging. For instance, it is implausible to prohibit logging if there is no real risk of harmful consequences such as a mudslide.⁶

The problem is not only that *NP* supports implausible verdicts in particular cases. The range of activities that will be prohibited is rather large. Hence, compliance with the principle against harm will have a stifling effect on the lives human beings lead. To appreciate how radical this is, imagine that everybody complies with the principle interpreted in these terms. Then nobody will have children anymore. So, if complied with, *NP* would basically be the end of humankind. Thus, the **STIFLING OBJECTION** reveals that the *NESS* Principle is too restrictive, as it prohibits too many actions.

⁶ I assume that these examples are such that one agent can indeed make a difference. However, there may be cases without a sharp boundary. It is commonly assumed that they are significantly more challenging. Yet, as Gunnemyr (2022) argues, the main theories of vagueness do not support this conclusion, as they postulate boundaries of their own. This suggests that solutions for cases with sharp boundaries generalize to cases without.

But why exactly is this? Many more actions are *NESS* conditions for harms in the prospective case as compared to the retrospective case, in a given time span. Consider an actual harm and an agent who contributed to it. Given that it occurred, there were enough other agents who contributed to it to form at least one minimally sufficient set. Consider next a possible harm and an agent who is now in a position to contribute to it. All that is required for their action to be a *NESS* condition for the harm is that there are enough agents who could possibly contribute to it so as to constitute a minimally sufficient set. It follows that a single action can be a *NESS* condition of many harms not even all of which are compossible.

It is important to distinguish the *STIFLING OBJECTION* from demandingness objections, to which I am not particularly sympathetic. The fact that a moral code is demanding is not as such an objection, as morality might just be very demanding (Goodin 2009). Furthermore, in certain times and places morality might be particularly demanding because wars, hunger and oppression are widespread. However, in the case at hand, many trivial actions will be prohibited for reasons that are not pressing at all. And the effects of those prohibitions are disproportional.

To find a solution, I go on to consider how *NESS* causation has been used in relation to backward-looking responsibility or blameworthiness. Matthew Braham and Martin van Hees (2012, 2018) have proposed what I call, for reasons to be discussed, ‘the Authorship Account’ of blameworthiness. I introduce it in section II.III. As I discuss in section II.IV, it provides the basis for a third version of the principle against harm, which I call ‘the Avoidance Principle.’

II.III. AUTHORSHIP AND BLAMEWORTHINESS

The Authorship Account of blameworthiness, proposed by Braham and van Hees (2012, 2018), can be reconstructed as follows (AA):⁷

A moral agent is blameworthy for a harm exactly if: (1) the action they performed was a *NESS* condition for it, (2) they performed it non-accidentally, and (3) they failed to maximize the avoidance potential without special justification.

⁷ As Braham and van Hees (2012, 605) explain, they are concerned with moral blameworthiness in a sense that resembles Watson’s (2004) notion of accountability and Scanlon’s (1998) notion of attributive responsibility.

The first condition, ‘the causal contribution condition’, reveals the role that *NESS* causation plays in the Authorship Account.⁸ The second condition is ‘the intentionality condition’.⁹ It is meant to ensure that the harm at issue bears the mark of ‘authorship’.¹⁰

Braham and van Hees argue that an agent can be blameworthy for making a causal contribution to a harm only if they had a reasonable opportunity to do otherwise. By this, they emphatically do not mean that the agent must have been in a position to prevent the harm. Instead, they should have been in a position to refrain from contributing to it. Furthermore, this must have been an eligible option. If it was not, they had a justification to make a contribution, which means they are not to blame for doing so. By way of example, suppose that, if one of the assassins in *ASSASSIN* had not taken the shot, their spouse and children would have been murdered. Under these circumstances, they could not reasonably have been expected to refrain from doing so, which means that this was not a reasonable or eligible option. Hence, had the assassin been in these circumstances, they would not be blameworthy.

Some cases are more complex. It might be, for instance, that an agent faced a choice between two options, both of which contributed to the harm in some way. Furthermore, the agent might have had more than two options. To handle such cases, Braham and van Hees (2012, 616) introduce the notion of an avoidance potential. This is the potential that an agent has to avoid authorship of a bad outcome. They formulate the relevant condition in terms of reasonability. Yet, they call it ‘the avoidance opportunity condition’ or ‘the avoidance condition’, which is why I formulated condition (3) in these terms (Braham and Hees 2012, 605; 2018, F98).

Thus, a moral agent is blameworthy for non-accidentally contributing to a harm if they could have performed an eligible action with a larger

⁸ To determine whether an action causes a harm requires a description of the situation. Such a description presupposes ‘a moral field’ (Braham and Hees 2012, 611–12). It is important to define it in a manner that is broad enough to include all candidates for blame.

⁹ The requirement of intentionality could be taken to imply that agents are not to blame for negligent or reckless actions. As this would not be particularly plausible, I leave this issue open by using the more flexible term ‘non-accidentally’.

¹⁰ Braham and van Hees also require that the agent is autonomous. Instead, I limit the analysis to moral agents. This is only a minor difference, however, as they equate autonomy with ‘normative competence’, which is the ability to make normative assessments (2012, 605–6). Elsewhere, they state that an autonomous agent “is able to act intentionally, to plan, and to distinguish right and wrong and good and bad” (Braham and Hees 2018, F98). Hence, their notion of ‘autonomy’ is similar if not identical to a common conception of moral agency (Wallace 1994).

avoidance potential. To escape blame altogether, they must have avoided authorship entirely. This in turn means that, to not be blameworthy at all, they must have maximized the avoidance potential of their actions (Braham and Hees 2018, F107n16). As will become important below, this notion is explicated in probabilistic terms: the agent must have maximized the probability of avoiding harm (Braham and Hees 2012, 617). As Braham and van Hees put it: you should do your “best to prevent an outcome by choosing a strategy that reduces the probability that your action is a *NESS* condition for an outcome” (2018, F99).

To illustrate the Authorship Account, consider JOYGUZZLING. It could be that the only alternative a joyguzzler had was to stay at home and play a game of scrabble with their family. However, perhaps they could also have written a letter to a representative in support of a higher energy tax on SUVs. The avoidance potential of such a political action is larger than that of playing scrabble (Cripps 2013; Hindriks 2023). In this second situation, refraining from joyguzzling is not sufficient for escaping blame. After all, they could have written the letter instead of playing scrabble. And this action has a larger avoidance potential. It follows that someone who played scrabble when it was an option to engage in political action is still an author of climate harm, at least to some extent. Hence, if the first two conditions are met as well, they are to blame for it, at least to some degree.

Imagine next that one of the assassins in ASSASSIN carried a laser pen in their gun bag. And suppose they could have used it to blind the other assassin, who would then have missed. In this case, the assassin could have pre-empted the assassination. As this would have minimized authorship or avoid it altogether, this is what they should have done if they are not to be held responsible. Thus, refraining from directly causing a harm is often not good enough. If you failed to maximize the avoidance potential of your actions, for instance by pre-empting it, you might still be to blame.

The Authorship Account pertains to backward-looking responsibility. Braham and van Hees do not consider whether it can be used to develop a plausible *NESS* account for forward-looking responsibility. This is what I go on to do.

II.IV. THE AVOIDANCE PRINCIPLE

Forward-looking responsibility becomes backward-looking responsibility as time goes by. More specifically, it does so at the last moment the

obligation could still have been fulfilled. This intuitive idea suggests that forward-looking and backward-looking responsibility are intimately related. More specifically, their agential, causal, and epistemic preconditions are basically the same. This is supported by the fact that, for an agent to be blameworthy in the sense at issue here, they must have performed an action that was wrong. In other words, they must have violated an obligation.

Now, suppose that the conditions under which agents are blameworthy for harm have been identified, then they can be used to construct an account of the conditions under which agents are obligated to refrain from performing the relevant actions. All one must do is formulate the account of blameworthiness in forward-looking terms. The only complication is that the agents might have an excuse. If they do, the action they performed was wrong, but they are not to blame for it (Wallace 1994). Factors that excuse are relevant only to backward-looking responsibility.¹¹ It follows that an account of backward-looking responsibility can be used to formulate an account of forward-looking responsibility.¹²

So, to formulate what I have called ‘the Avoidance Principle’, I reformulate the conditions of the **AUTHORSHIP ACCOUNT** in forward-looking terms (*AP*):

AP There is a *pro tanto* obligation not to perform an action that can be a *NESS* condition for harm if it does not maximize the avoidance potential.

The first part of *AP* mirrors the causal condition. The second part, the avoidance condition. Because it specifies a *pro tanto* obligation, *AP* accommodates the role that special justifications play in this condition of the Authorship Account. It should be taken to imply that, for an action to be permissible, it must be the eligible action with the largest avoidance potential. Furthermore, the notion of an avoidance potential is to be

¹¹ The fact that an agent might have a justification can be accommodated by postulating a *pro tanto* obligation.

¹² Van de Poel conceives of the relation between the two as follows: “If A has the obligation to see to it that some state of affairs X is the case and X happens not to be the case, A is accountable [answerable] for ‘not X’. If A is not able to give a satisfactory account for ‘not X’, A is also blameworthy” (2011, 44). This supports the idea that an account of backward-looking responsibility can be used to formulate an account of forward-looking responsibility.

understood in probabilistic terms. In principle, an agent is to perform actions that minimize the probability of harm.¹³

How plausible is this? Consider the assassin with a laser pen in their gun bag. *AP* requires them to use it and blind the other assassin such that they miss their shot. Similarly, the potential joyguzzler who is able to write a letter should in principle do so. It may well be that morality does indeed require this. However, *AP* also supports verdicts that are less plausible. Consider someone who wants to go to a club. Presumably, going for a walk is less likely to result in harm. Absent other considerations, this is what they should do. This reveals that minimizing the probability of harm is a rather strong requirement. Consider fishing next. Suppose that there is an extremely small probability that the species concerned goes extinct. There might be activities you could engage in that are even less likely to result in harm. If so, you should in principle refrain from fishing. This line of reasoning generalizes to the logging example mentioned earlier. And it suggests that the STIFLING OBJECTION applies not only to *NP*, but also to *AP*. In this case, the problem is that *AP* merely considers the probability of no harm, which must be maximized. The probability of harm plays no role in it.

In response, one might say that these options are not eligible. It is too much to ask for people to refrain from clubbing, fishing, or logging. Earlier I suggested that an action is ineligible exactly if the agent has a special justification not to perform it. This in turn means that they have an overriding obligation. However, this is an incidental matter. It does not provide a systematic solution to the problem. Hence, the STIFLING OBJECTION applies to *AP* as well. Perhaps it can be answered by explicating the notion of eligibility in a different manner. However, if this notion is understood in a more liberal manner, the requirement that the probability of harm be minimized might become trivial. So, it remains to be seen whether *AP* can be saved by amending it in some such manner.

How, if at all, does this criticism of *AP* reflect on the Authorship Account of blameworthiness? Earlier I suggested that the only real difference between forward-looking and backward-looking responsibility is time. If this is indeed the case, an account of the one is inextricably tied to the other. This means that, if *AP* is too restrictive, then so is the Authorship Account. Now, the STIFLING OBJECTION reveals that *AP* does indeed prohibit too many actions. Given the assumption mentioned, it follows

¹³ *AP* does not require that the action be performed non-accidentally. Presumably, whether it does bears only on blameworthiness and not on impermissibility.

that the Authorship Account of blameworthiness is also too restrictive. More precisely, there will be too many situations where the agent must justify the fact that they performed an action that is in principle impermissible.

However, my aim here is not to evaluate the Authorship Account. Instead, it is to identify a plausible formulation of the moral principle against harm. In section I, I have argued that a moral principle that prohibits only difference-makers is too permissive. In this section, I have argued that moral principles that are based on *NESS* causation, in particular *NP* and *AP*, are too strict, as they are susceptible to the *STIFLING OBJECTION*. This suggests that the way forward is to find an account of causation that is stronger than *NESS* and weaker than difference-making. Security dependence fits the bill (Gunnemyr and Touborg 2023a, 2023b). So, I go on to discuss whether it solves the Problem of Collective Harm.

III. SECURITY DEPENDENCE

III.I. THE SECURITY PRINCIPLE

According to Mattias Gunnemyr and Caroline Touborg (2023a, 2023b), causation is to be understood in terms of security dependence. This means that causes make their effects more secure. If reformulated in these terms, the principle against harm prohibits actions that make harm more secure.

Security dependence is premised on the idea that causation is contrastive: an event causes one event (*E*) rather than another one (*E**), which is incompatible with it. In the nearest world where *C* does not occur, *E* is less secure and *E** is more secure. It follows that a cause makes it less easy for its effect to fail to materialize. To what extent depends on the distance between the actual world and the nearest possible world in \hat{H} where it does not occur. The wider the range of circumstances under which it occurs, the more secure it is. As I discuss in more detail below, this is to be evaluated relative to a ‘relevant possibility horizon’ \hat{H} , a range of possible worlds that are pertinent to the issue at hand.¹⁴

¹⁴ The world in which *C* does not occur should be substantially different, so as to avoid that, in this world, an event that is rather similar to *C* causes *E* (Gunnemyr and Touborg 2023b, 14) Furthermore, if a harm occurs in every possible world, it is infinitely secure both in the actual world and in the closest world where a candidate cause does not occur. In that case, it is not a cause after all (Gunnemyr and Touborg 2023b, 19).

Gunnemyr and Touborg (2023a, 2023b) also require that an effect is connected to its cause by a process that consists of a time SEQUENCE OF CONTRIBUTIONS TO THE HARM (*SC*):

SC *C* is a cause of *E* rather than *E** within possibility horizon \hat{H} just in case: (a) *C* is process-connected to *E*; and (b) there is at least one world in \hat{H} where *C* does not occur, and in the closest-to-@-at-*t* world(s) in \hat{H} where *C* does not occur, *E* is less secure at *t* and *E** is more secure at *t* than they are in @ (Gunnemyr and Touborg 2023b, 15-17).

Gunnemyr and Touborg (2023a, 344) give a few guidelines as to how to specify \hat{H} . The first step is to identify the agents in the relevant situation and their options. Each combination of choices, each choice-assignment, is relevant. Furthermore, \hat{H} can usually be restricted to worlds in which the non-agential features are the same as in the actual world.

These rules have consequences as to which events qualify as causes. To get a sense of this, imagine that your favorite club has a slow night. Instead of the usual five hundred, there are now only five people. In this situation, the possibility that anyone faints due to overcrowding is not relevant. Because of this, it does not feature in the relevant possibility horizon \hat{H} . This implies that you entering the club does not make someone's fainting more secure. Suppose next that a particular fish species is about to go extinct. It has been overfished for years. And there is no way back. In that case, the possibility that the species survives is so remote that it should not be included in \hat{H} . It follows that you catching a fish does not cause it to go extinct.¹⁵

Gunnemyr and Touborg (2023a) use their account of security dependence to shed light on the reasons people have for contributing to a harm. They argue that there is reason to refrain from an action for reasons of harm only if the action makes a harmful outcome more secure. Elsewhere, they use it to identify the conditions under which an agent is blameworthy for performing an action that is a security cause of harm (Gunnemyr and

¹⁵ The notion of a possibility horizon is similar to that of a moral field in Braham and van Hees (2012). However, as mentioned in footnote 8, they argue that it is important to define it in a manner that is broad enough. Here it becomes apparent that the relevant possibility horizon can be specified in a way that is too broad. If too many events are included, too many of them will qualify as causes.

Touborg 2023b). On the assumption that blameworthiness presupposes wrongdoing, this supports what I call the **SECURITY PRINCIPLE (SP)**:

SP There is a *pro tanto* obligation not to perform an action that makes a harm more secure.

By way of example, consider John, who is disinclined to help other people (*Indifferent JOHN*). When he sees a child struggling in the water, he continues to walk along the beach. The child drowns, even though John could easily have saved her. The first question to ask is whether his inaction caused the child's death. Consider the nearest world in which John is not indifferent but inclined to help people, w_1 . In w_1 , John jumps into the water and saves the child. Furthermore, the child's drowning is less secure in w_1 as compared to @. Finally, the child's survival is more secure in w_1 . Hence, John's inaction does indeed cause the child's death.¹⁶

Security dependence is indeed intermediate between *NESS* causation and difference-making. An action that makes an outcome more secure will be a *NESS* condition for it. However, an action can make a contribution to an event without making it more secure. It might be, for instance, that the possibility that the event fails to occur is so distant as to be irrelevant. As mentioned above, the possibility of overcrowding is irrelevant when a club is having a slow night.¹⁷ Furthermore, all difference-makers make their effects more secure. Yet, not all events that make other events more secure are difference-makers. This is confirmed by the claim that a cause makes a difference to the security of the effect, but not necessarily to the effect as such (Gunnemyr and Touborg 2023a, 343).

Given that security causation is indeed intermediate between *NESS* causation and difference-making, it might solve the STIFLING OBJECTION.

III.II. THE STIFLING OBJECTION

According to the STIFLING OBJECTION, discussed in sections II.II and II.IV, *NP* and *AP* prohibit too many actions that contribute to collective harms. In fact, it supports universal or near universal prohibitions of activities such as clubbing and deforestation, irrespective of whether there is a genuine risk of harm. This is problematic because many such activities seem to be

¹⁶ Gunnemyr and Touborg (2023b, 21–22) also argue that John is to blame for what he did, which presupposes that it was wrong.

¹⁷ Another reason why security causation is more demanding is that, in contrast to *NESS* causation, it includes a process condition (condition b).

permissible, at least up to a point. The question at issue here is whether the STIFLING OBJECTION also applies to *SP*.

Not every contribution to a collective harm makes it more secure. There has to be a relevant possibility of harm. Consider eating a fish. If the possibility of its species' extinction is not part of the relevant possibility horizon, then doing so does not make extinction more secure. In that case, catching a fish is not a cause of the extinction of the species. Furthermore, doing so will not be prohibited, at least not for reasons of harm. This suggests that the security account does indeed have an answer to the STIFLING OBJECTION.

This answer to the objection generalizes to other cases. Exactly how depends on the intuitions you have, as they determine which worlds feature in \hat{H} . In this respect, *SP* is pretty flexible. This is both an advantage and a disadvantage. It makes the proposal attractive for people with different intuitions. Because of this, it will confirm whatever intuitions you have about which prohibitions qualify as stifling. The downside is that this makes the proposal relatively subjective. So, it would be good if there were a more objective way to arrive at verdicts.

The upshot is that, in contrast to *NP* and *AP*, *SP* has a pretty convincing answer to the STIFLING OBJECTION. The reason for this is that it treats certain risks as too remote to be relevant for who causes what. It follows that actions that contribute to a harm that is improbable because it is too remote are not prohibited for reasons of harm. In section 4, I argue that my own proposal, the threshold probability account, is more objective than *SP*. But first I discuss a second problem from which *SP* does suffer, which arises when not enough others are willing to comply with the principle against harm.

III.III. THE PROBLEM OF UNWILLINGNESS

Alice, Bill, and Cecil live around a lake. Each paints their boat. Due to the solvent used, the water is polluted such that the fish in the lake die. Crucially, they would have died even if only two individuals had painted their boats using this solvent. So, at least two of them must stop using it if they are to save the fish. Björnsson (2011, 2014), who proposed this example, points out that, intuitively, each of the three lakers should stop using the solvent. *SP* supports this verdict. Alice, Bill, and Cecil each make the death of the fish more secure. This means that, unless someone has a special justification, each of them ought to stop using the solvent (Gunnemyr and Touborg 2023a, sec. 4; 2023b, sec. 5).

Imagine next that none of the lakers are willing to stop using the toxic solvent. In fact, they are ‘robustly unwilling’. By this, I mean that there is nothing anyone could do to change their minds. Each of them is utterly indifferent about the fate of the fish. And they only care about doing things the way they want them to do, including using the solvent of their choice. And they care a lot about this. As a consequence, their behavior remains the same across a wide range of circumstances. I call this version of the example ‘LAKE’.¹⁸

For something to be feasible, it must be possible in the relevant circumstances (Wiens 2015). So, is it feasible for Alice, Bill, and Cecil to save the fish? These lakers are all robustly unwilling to change their behavior. Because of this, it is impossible to do so in the circumstances at issue. Consider Alice. Even if she were to change her behavior, Bill and Cecil would not. Hence, the fish would still die. The same holds for Bill as well as for Cecil. This is true even if one tries to convince the others. Thus, those living around the lake are constrained by each other. Given these circumstances, it is not feasible to save the fish.

Because of this feature, LAKE illustrates what I call ‘the problem of unwillingness’. The question is whether, in these circumstances, the three lakers have an obligation to refrain from using the solvent. I propose they do not. None of them is obligated to change their behavior if all of them are robustly unwilling. To be sure, had there even been a remote possibility that the lakers would have changed their mind, then each might have been obligated to stop using the solvent. However, in the situation at issue, saving the fish is not a real option. Thus, the problem of unwillingness reveals that *harm cannot obligate if avoiding it is not feasible*.

This is why LAKE is such an interesting test case: although it is possible to save the fish, doing so is not feasible. The scope of *NP* depends on what is possible, not on what is feasible. Hence, *NP* requires the lakers to change their behavior. This verdict is misguided, I propose, because the probability that the fish will be saved is zero. But what about *SP*? Just as *NP*, this principle is not sensitive to the fact that the lakers are robustly unwilling to change their behavior. This is due to the fact that their unwillingness does not affect the relevant possibility horizon \hat{H} . Hence, it is not morally relevant. It follows that, according to *SP*, the three lakers have a *pro tanto* obligation to stop using the solvent. However, as I have just argued, this seems to be the wrong verdict. Hence, *SP* is too restrictive.

¹⁸ This version is similar to the one Björnsson (2014, 108) calls ‘Known Reluctance’, except for the fact that the term ‘reluctance’ is weaker than ‘robust unwillingness’.

Can this problem be solved? Gunnemyr and Touborg (2023a, 356–58) claim that intuitions vacillate in this case. They thereby recognize that some will regard the verdict that *SP* supports as mistaken. At the same time, they argue that their account can be amended so as to support the verdict that the lakers are not obligated to do anything. The thing to do is to specify \hat{H} , not in the broad manner that they prefer, but in a narrow manner. This means that \hat{H} is restricted to worlds where nobody but the agent under consideration changes their behavior. When adjusted in this way, the proposal does not suffer from the problem of unwillingness.¹⁹

However, the flexibility of *SP* is both a blessing and a curse. Due to its flexibility, *SP* can be amended to support (what I take to be) the correct verdict. At the same time, the fact that it is so flexible also implies some arbitrariness. This strengthens the claim made in section III.II that it would be good if there were a more objective way to arrive at verdicts. In light of this, I go on to formulate an alternative account, one that attributes a more central role to probabilities.²⁰

IV. THE THRESHOLD PROBABILITY ACCOUNT

IV.I. THE PROBABILITY OF MAKING A DIFFERENCE

The threshold probability account revolves around the probability that an action makes a difference to a harm. And it supports what I call the **THRESHOLD PRINCIPLE'** (*TP*):

TP There is a *pro tanto* obligation not to perform an action exactly if the probability that it brings about a harm is too high.

As before, an action that brings about harm makes a difference to it. In light of this, I call the probability that features in *TP* 'the difference-making probability'. If it is not low enough, it exceeds a certain threshold. In that case, there is a *pro tanto* obligation not to perform it. Conversely, it

¹⁹ Gunnemyr and Touborg propose that \hat{H} should not include a possibility that "someone might do something that is not an option for her" (2023a, 344). Presumably an action is not an option for an agent if they are robustly unwilling to perform it. This suggests that \hat{H} should have been specified in a narrow manner all along. Yet, they also claim that \hat{H} should include all possible choice assignments, which implies the opposite (Gunnemyr and Touborg 2023a, 344).

²⁰ As I have just argued, events with zero probability are irrelevant to who is obligated to do what. Another reason for invoking probabilities is to do justice to cases in which there are no causally sufficient conditions for outcomes. Björnsson (2011, 185) criticizes *NESS* causation for its inability to do accommodate such probabilistic cases.

is permissible to contribute to a harm if this probability is so low as to be morally acceptable (Hindriks 2022).

To give more content to this proposal, I need to say more about the probability of making a difference. As discussed in section I, an action that makes a difference to an outcome is individually necessary and jointly sufficient for it in the circumstances at hand. In light of this, I propose that the probability of making a difference is complex. More specifically, it is to be analyzed in terms of two conditional probabilities, the probability of harm conditional on an action and the probability of no harm conditional on its omission.

Let a_i be an action of agent i that, if performed, makes a contribution to a harm (H) in situation S . The probability of harm conditional on a_i given the actions of the others is: $P(H|a_i, a_j...a_n, S)$. The probability of no harm conditional on agent i not performing the action given what the others do is: $P(\sim H|\sim a_i, a_j...a_n, S)$. Together, these probabilities determine the probability of making a difference. Let ' d ' be the number of contributions required for bringing about some collective harm. The closer this is to the actual number of contributors, the higher the probability of making a difference.

Now, suppose that $n = d$. Then there are neither too few nor too many contributions for bringing about the collective harm: each action is individually necessary and jointly sufficient. It follows that the probability of making a difference equals 1. The same holds for the two constituent probabilities: $P(H|a_i, a_j...a_n, S) = 1$ and $P(\sim H|\sim a_i, a_j...a_n, S) = 1$. The conditional probability of harm can be interpreted as the probability that the harm is underdetermined. And the conditional probability of no harm interpreted as the probability that it is overdetermined. If one of them equals zero, the difference-making probability equals 0 as well.

Because the probability of making a difference is complex in the way explained, the threshold probability account features two threshold requirements (TR). They pertain to the (conditional) probability of harm and the (conditional) probability of no harm:

TR1 An action is impermissible for reasons of harm only if the probability of harm is too high.

TR2 An action is impermissible for reasons of harm only if the probability of no harm is too high.

To see whether these requirements are met, the actual probabilities must be compared to two threshold probabilities. These are the threshold probability of underdetermination (P_{tu}) and the threshold probability of overdetermination (P_{to}).

According to the threshold probability account, an action is in principle impermissible for reasons of harm exactly if both thresholds are met. This is captured by the **THRESHOLD PRINCIPLE (TP)**:²¹

TP There is a *pro tanto* obligation not to perform action a_i exactly if

- (i) $P(H|a_i, a_j \dots a_n, S) < P_{tu}$ and
- (ii) $P(\sim H|\sim a_i, a_j \dots a_n, S) < P_{to}$

If either one of these conditions is not met, then the action is permitted, or at least not prohibited for reasons of harm.²²

Just as *SP*, *TP* is intermediate between *DP* and *NP*. Instead of difference-making, it invokes the probability of making a difference. This means that it is more restrictive than *DP*. Furthermore, it prohibits actions that are *NESS* conditions for harm only if the probability that they make a difference is significant. Hence, *TP* is more permissive than *NP*. I go on to argue that *TP* solves both the **STIFLING OBJECTION** and the **Unwillingness Objection**. The former implies that it is also more permissive than *AP*; the latter that it is less restrictive than *SP*.

IV.II. THE THRESHOLD PROBABILITIES

Earlier I proposed that an action is impermissible for reasons of harm only if the probability that it makes a difference is too high. This can be made more precise in terms of *TP*. It means that the harm probability and the no harm probability exceed the corresponding threshold probabilities. But how high are they? The first thing to note is that they can in fact be low, if not extremely low. And if they are, an action can be impermissible even though it is rather unlikely that the harm would materialize if it were performed.

²¹ I have presented *TP* as an alternative to *DP*. However, it could also be interpreted as practical guidance as to how to apply *DP*. In that case, it captures what Burri (2022) calls ‘a duty of care’.

²² An earlier version of this proposal, which I called ‘the prospect account’, features condition (i) but not condition (ii). Instead, it invokes the extent to which an individual agent affects the probability of harm (Hindriks 2022). The change is due to the analysis of the notion of making a difference presented earlier.

The threshold probabilities depend on the size of the harm. The greater the harm, the lower they are. According to some normative theories, they will also depend on the opportunity costs. These are the sacrifices agents would make by showing restraint and performing the next best action. If this is indeed the case, they increase monotonically with the opportunity costs. Hence, the threshold probabilities are particularly low if the outcome is very harmful and the opportunity costs are very low.

JOYGUZZLING may well be such a case. Presumably, the climate harm at issue is rather large. And the pleasure someone derives from playing scrabble need not be much lower than that of going for a drive. If the difference is small, the opportunity cost is low. This suggests that the threshold probabilities are low, if not extremely low. So, *TP* might actually prohibit this activity. If it does, joyguzzling exposes others to a risk of harm that is morally unacceptable. Yet, it could also be that the difference-making probability is so small as to be negligible in this case.

The threshold probability account provides a structure that can be made more precise by combining it with a particular normative theory. For instance, a rule consequentialist version will assume that the threshold probabilities depend on total amounts: H and C . In contrast, a contractualist version will invoke per capita amounts instead: h and c . Furthermore, a rule consequentialist version will give equal weight to harms and opportunity costs, while a contractualist version need not do so. If the threshold account is consistent with Kantianism, H is presumably all that matters. Finally, the account does not require maximizing the value of the consequences of an action. Hence, it is incompatible with act consequentialism.²³

IV.III. THE PROBLEM OF UNWILLINGNESS

Does the threshold probability account solve the problem of unwillingness? Recall that, in LAKE, the death of the fish is overdetermined. Three individuals use the solvent, while two suffice for the harm. Hence, $P(H|a_1, a_2 \dots a_m, S) = 1$. Furthermore, the lakiers are robustly unwilling to change their behavior. Because of this, it is not feasible to prevent the harm, which means that $P(\sim H|\sim a_1, a_2 \dots a_m, S) = 0$. This means that the probability of making a difference is 0. This in spite of the fact that the possibility is

²³ So, it is inconsistent with consequentialist solutions to the problem of collective harm such as Singer (1980), Norcross (2004) and Kagan (2011). See Hindriks (2022) for more on this.

close. It follows that *TP* supports the intuitive verdict in LAKE: insofar as reasons of harm are concerned, using the solvent is permissible.

But how can this be given that the behavior of the lakers issues in a harmful outcome? For one thing, their contributions are harmless. Furthermore, it is not even feasible for them to prevent the harm. To be sure, it is in principle possible to avert the collective harm. However, this is merely of theoretical interest and, as such, irrelevant for practical purposes.²⁴

Gunnemyr and Touborg disagree with this because they subscribe to the principle of moral harmony: “The principle states that if all the agents involved in a situation act as they have objective reason to act, then the resulting pattern of behavior will lead to the best attainable outcome” (2023a, 336).²⁵ This implies that, in situations that give rise to suboptimal outcomes, at least one individual failed to act as they had objective reason to do. In LAKE, this means that if the fish die then at least one of the lakers did not act as they had reason to do. Furthermore, as they are placed symmetrically, they all had this reason.

On the intended reason, ‘the best attainable’ is equivalent to ‘the best that is possible.’ But on this interpretation, the principle of moral harmony is implausible. As I argued in section III.III, harm cannot obligate if avoiding it is not feasible. To be sure, a moral ideal can ground obligations even if it takes several steps to get there. It might be, for instance, that people have an obligation to protest in order to get those in power to take action. In a similar vein, it has been argued that they can have an obligation to moralize behavior, mobilize others, or form a collective agent (Collins 2013; Lawford-Smith 2015; Hindriks 2019). But if there are such obligations, this means that the relevant outcomes are feasible. The reason for this is that a moral ideal does not license ignoring the bounds of feasibility. The upshot is that *TP* solves the problem of unwillingness in a plausible manner.

IV.IV. THE STIFLING OBJECTION

According to the STIFLING OBJECTION, a principle that is supposed to solve the problem of collective harm is implausible if it is so restrictive that life

²⁴ Björnsson (2014) agrees that none of the lakers has an obligation to do something. At the same time, he argues that the lakers have an irreducibly collective or ‘essentially shared’ obligation to stop polluting the lake. In this way, he tries to do justice to the fact that prevention is not feasible and to the fact that it is possible.

²⁵ Gunnemyr and Touborg (2023a, 341n11) invoke the principle of harmony to criticize an earlier version of the threshold probability account, mentioned in footnote 22.

as we know it would basically come to a halt if people were to comply with it. This objection counts against *NP* and *AP* (section II). But *SP* has an answer to the objection (section III.II). Here I argue that *TP* does so as well. As discussed in section IV.II, the threshold probabilities increase monotonically with the opportunity costs. Because of this, those costs put a limit on how many actions *TP* prohibits.

As discussed, the opportunity cost of joyguzzling is in all likelihood pretty low. However, the opportunity cost of driving an SUV on a Sunday afternoon can be high. Braham and van Hees (2012, 617) consider someone who needs to get to the hospital urgently because their wife is in labor. It may be that the only way to get there fast enough is by means of the SUV. This would mean that the trip to the hospital is permissible. It follows that actions that are in principle prohibited might be permitted if they are particularly costly for the agent. As I see it, this counts in favor of the threshold probability account, especially in comparison to *SP*. Including opportunity cost adds further structure to the proposal, which makes it more objective.

SP can be made more or less restrictive by adjusting the possibility horizon. As discussed, it can even be amended to land on the other side of the fence insofar as the Problem of Unwillingness is concerned. This enables it to accommodate different intuitions. However, it also makes the verdicts it supports somewhat arbitrary. In this respect, *TP* rules out more options (without being too rigid). Furthermore, it is significantly simpler. The upshot is that both accounts have an answer to the STIFLING OBJECTION. Yet, *TP* is more plausible than *SP*.

V. CONCLUSION

Collective harms are due to several individual actions that are as such harmless. The Problem of Collective Harm is the challenge of determining whether and when it is impermissible to contribute to such a harm. One option is to regard only those actions that make a difference to the outcome as impermissible. Another option revolves around *NESS* causation. It says that in principle any action that is a necessary element of a set of actions that is sufficient for the outcome is prohibited for reasons of harm. I have argued that the former is too permissive, while the latter is too restrictive (the STIFLING OBJECTION). So, what is needed is an intermediate account.

One option is to formulate an account that is based on *NESS* causation, but less restrictive. The proposal I considered prohibits making a causal

contribution to a harm only if the action does not maximize its avoidance potential. As it turns out, this account does not provide an adequate answer to the STIFLING OBJECTION. Another option is to invoke yet another conception of causation, secure causation. This security account requires action only if it makes the harm less secure. But it is too restrictive in another manner, as it does not recognize the moral relevance of robust unwillingness. Hence, it suffers from the Problem of Unwillingness.

The threshold probability account revolves instead around the probability of making a difference. It is permissible to contribute to a harm only if this probability is too high. However, it is zero if everyone is robustly unwilling to refrain from doing so. In that case, it is not feasible to prevent the harm, which means that no one is required to do anything. In this way, the account solves the Problem of Unwillingness. To answer the STIFLING OBJECTION, I observed that the probability of making a difference is often negligible. And I argued that it is permissible to contribute to a harm if its possibility is so remote that it need not be reckoned with.

Thus, the threshold probability account handles both possibilities that are remote and improbable and possibilities that are close but improbable in a plausible manner. I conclude that this probabilistic account is to be preferred to the three causal proposals.

REFERENCES

- Björnsson, Gunnar. 2011. "Joint Responsibility Without Individual Control: Applying the Explanation Hypothesis." In *Moral Responsibility*, edited by Jeroen van den Hoven, Ibo van de Poel, and Nicole Vincent, 181–199. Dordrecht: Springer.
- Björnsson, Gunnar. 2014. "Essentially Shared Obligations." *Midwest Studies in Philosophy* 38 (1): 103–120.
- Braham, Matthew, and Martin van Hees. 2009. "Degrees of Causation." *Erkenntnis* 71 (3): 323–344.
- Braham, Matthew, and Martin van Hees. 2012. "An Anatomy of Moral Responsibility." *Mind* 121 (483): 601–634.
- Braham, Matthew, and Martin van Hees. 2018. "Voids or Fragmentation: Moral Responsibility For Collective Outcomes." *The Economic Journal* 128 (612): 95–113.
- Burri, Susanne. 2022. "Conceptualising Morally Permissible Risk Imposition without Quantified Individual Risks." *Synthese* 200 (5): 1–22.
- Collins, Stephanie. 2013. "Collectives' Duties and Collectivization Duties." *Australasian Journal of Philosophy* 91 (2): 231–248.
- Cripps, Elizabeth. 2013. *Climate Change and the Moral Agent*. Oxford: Oxford University Press.
- Goodin, Robert E. 2009. "Demandingness as a Virtue." *The Journal of Ethics* 13 (1): 1–13.
- Gunnemyr, Mattias. 2022. "Making a Vague Difference: Kagan, Nefsky and the Sorites Paradox." *Inquiry*, 1–26.

- Gunnemyr, Mattias, and Caroline Torpe Touborg. 2023a. "Reasons for Action: Making a Difference to the Security of Outcomes." *Philosophical Studies* 180 (1): 333-362.
- Gunnemyr, Mattias, and Caroline Torpe Touborg. 2023b. "You Just Didn't Care Enough: Quality of Will, Causation, and Blameworthiness for Actions, Omissions, and Outcomes." *Journal of Ethics and Social Philosophy* 24 (1): 1-35.
- Halpern, Joseph Y. 2008. "Defaults and Normality in Causal Structures."
- Hindriks, Frank. 2019. "The Duty to Join Forces: When Individuals Lack Control." *Monist* 102: 204-220.
- Hindriks, Frank. 2022. "The Problem of Insignificant Hands." *Philosophical Studies* 179: 829-854.
- Hindriks, Frank. 2023. "When to Start Saving the Planet?" *Journal of Ethics and Social Philosophy* 23 (3): 397-419.
- Kagan, Shelly. 2011. "Do I Make a Difference?" *Philosophy and Public Affairs* 39 (2): 105-141.
- Lawford-Smith, Holly. 2015. "Unethical Consumption and Obligations to Signal." *Ethics & International Affairs* 29 (3): 315-330.
- Lewis, David K. 1973. *Counterfactuals*. Oxford: Blackwell.
- List, Christian, and Kai Spiekermann. 2013. "Methodological Individualism and Holism in Political Science: A Reconciliation." *American Political Science Review* 107 (4): 629-643.
- Nefsky, Julia. 2011. "Consequentialism and the Problem of Collective Harm: A Reply to Kagan." *Philosophy and Public Affairs* 39 (4): 364-395.
- Nefsky, Julia. 2017. "How You Can Help, without Making a Difference." *Philosophical Studies* 174: 2743-2767.
- Nefsky, Julia. 2019. "Collective Harm and the Inefficacy Problem." *Philosophy Compass* 14 (4): 1-17.
- Norcross, Alistair. 2004. "Puppies, Pigs, and People: Eating Meat and Marginal Cases." *Philosophical Perspectives* 18: 229-245.
- Poel, Ibo van de. 2011. "The Relation Between Forward-Looking and Backward-Looking Responsibility." In *Moral Responsibility* edited by Jeroen van den Hoven, Ibo van de Poel, and Nicole Vincent, 37-52. Dordrecht: Springer.
- Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Singer, Peter. 1980. "Utilitarianism and Vegetarianism." *Philosophy & Public Affairs* 4 (9): 325-337.
- Sinnott-Armstrong, Walter. 2005. "It's Not My Fault: Global Warming and Individual Moral Obligations." In *Perspectives on Climate Change: Science, Economics, Politics, Ethics (Advances in the Economics of Environmental Resources, Vol. 5)*, 285-307. Leeds: Emerald.
- Thompson, Dennis. 1980. "Moral Responsibility of Public Officials: The Problem of Many Hands." *American Political Science Review* 74 (4): 905-916.
- Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge (MA): Harvard University Press.
- Watson, Gary. 2004. *Agency and Answerability*. Oxford: Oxford University Press.
- Wiens, David. 2015. "Political Ideals and the Feasibility Frontier." *Economics and Philosophy* 31 (3): 447-477.

- Wright, Richard W. 1985. "Causation in Tort Law." *California Law Review* 73 (6): 1735-1828.
- Wright, Richard W. 1988. "Causation, Responsibility, Risk, Probability, Naked Statistics, and Proof." *Iowa Law Review* 73: 1001-1077.
- Wright, Richard W. 2011. "The NESS Account of Natural Causation: A Response to Criticisms." In *Perspectives on Causation*, edited by Richard Goldberg, 285-322. Oxford: Hart Publishing.

Frank Hindriks is professor of Ethics, Social and Political Philosophy at the University of Groningen. He was Assistant Director at the Erasmus Institute of Philosophy and Economics (EIPE), founding member of the International Social Ontology Society and founding editor of the *Journal of Social Ontology*. He works primarily on social ontology and collective ethics. And he has published in a wide range of journals, including *Philosophical Quarterly*, *Philosophical Studies* and the *Journal of Institutional Economics*.

Contact e-mail: <f.a.hindriks@rug.nl>