Outcome Responsibility: Fallible Beings Acting in an Uncertain World

Alexa Zellentin

School of Politics & International Relations, UCD

Outcome Responsibility: Fallible Beings Acting in an Uncertain World

Alexa Zellentin

Abstract

With regard to historical emissions contributing to climate change we find ourselves in a situation where agents blamelessly caused a problem for uninvolved others. When arguing that the originators of the problem should be considered responsible for rectifying it, we face a number of problems. One of these problems is that assigning liability to blameless originators is usually justified with reference to social practices (e.g. strict liability), which are deemed as beneficial to all participants. Such practices, however, need to be in place before the problematic act came to pass. It would be unfair to retrospectively argue that such a rule would have been convenient and we should therefore act as if it had been in place where the agents affected where not yet aware of such a rule and its demands on them. On the other hand, justifying rectificatory duties with reference to the responsibility of having caused the problem alone, requires a kind of morally significant responsibility the preconditions for which might not be satisfied in cases where fallible agents act in an uncertain world. This paper takes up Honoré's and Miller's account of outcome responsibility, which establishes duty generating responsibility in cases where the conditions for blameworthiness are not fulfilled. In particular, this paper analyses the responsibility of agents whose capacities sometimes inexplicably fail as well as responsibility in cases where agents cannot know whether the safeguards they put in place are sufficient to prevent harming others. It addresses some important critiques of the original accounts of outcome responsibility and shows how the open questions can be resolved with regard to Raz's conception of a sphere of secure competence. The main conceptual idea defended in this paper is that the agency of fallible beings in an uncertain world is genuine agency and thus can be the source of rectificatory duties. The main normative point is to suggest that we ought to think ahead which social practices best balance incentives for innovation with a regime that adequately ensures that the victims of the unintended side effects of fallible agents acting in an uncertain world are not left to deal with the outfall unaided.

Alexa Zellentin is a faculty member at the School of Politics and International Relation (SPIRe) at University College Dublin. Areas of interest: normative political theory, esp. climate justice, rectificatory justice, liberal neutrality, and justice in education. Her monograph on liberal neutrality is published in the series Ideen und Argumente (De Gruyter), journal articles appear e.g. in the Journal of Applied Philosophy, Res Publica, and Climatic Change. She is a board member of the editorial committee of Ethical Theory and Moral Practice.

Introduction

We know situations where we enter a shop with breakable goods, are as careful as we can be and yet somehow manage to break something. Often we are expected to pay for the broken goods and usually we apologise and pay without making a fuss.[1] We can also easily imagine setting up a bonfire in the garden, taking all our usual precautions and expecting that our usual level of care will once more be sufficient to prevent accidents, and yet somehow this time we manage to set the neighbour's shed on fire. We know that there are rare cases where despite all reasonable precautions an unexpected combination of a strong gust of wind and dry piece of wood sends off sparks sufficiently strong to cause trouble. Usually, there is no question but to apologise for the mishap and compensate the neighbour for the damage.[2] In legal contexts, this would be a matter of 'strict liability'. The default position of 'the causer pays' seems fairer than the only alternative of 'victim pays'. While we often apologise and compensate in such cases, we do not feel that it would be appropriate to blame us for the damage (provided we took all reasonable precautions) – these were genuine accidents outside our control.[3] We are fallible agents acting in an uncertain world, and there only is a certain level of foresight and care that can be expected of us. This paper analyses these kinds of cases where we have good reasons to consider the conditions of genuine agency as fulfilled even though the conditions of blameworthiness are not.

Both the apology and the compensation in the example are the result of social conventions. In Robin Zhang's understanding any responsibility here is a case of *responsibility as accountability* but has nothing to do with *responsibility as attributability*, which is concerned with actions (and outcomes) that can be seen as "manifestations of our ends, commitments, or values".[4] This paper will not dispute this distinction, however, it will explore the idea that responsibility as accountability does not necessarily presuppose actually existing social practices (though this often is the case) but rather applies in all cases where it is "appropriate for others to hold us accountable"[5] for certain outcomes. There are two ways to justify why we should introduce or continue social practices. One way is to refer to the usefulness of the solution. We know that accidents happen and that it is extremely difficult to identify whether someone really took all reasonable precautions. To avoid invasive investigations, it is reasonable – where there are no clear indications to the contrary – to assume no ill intent but nonetheless ask the person who caused the accident to take responsibility for the damage. To avoid hardship, we add social practices like offering liability insurance for the cases where we mess up and accidentally harm others and home insurance for cases where others break our

---

[1] The exception are cases where shops were set up in a disaster-prone way and we feel the responsibility should lie with those who balance breakable goods so precariously.

[2] An exception might be a case where the neighbour – unknown to us – coated the shed in something flammable so that even the tiniest spark that under normal circumstances would not have caused even a mark suffices to set the whole shed alight. Once more, here it might be reasonable to argue that the responsibility for the damage should lie with the agents creating a dangerous environment.

[3] These examples are modelled after two key examples that Miller uses to explain his understanding of outcome responsibility. Why I used a variation to the original cut glass shop example, namely an ordinary person as opposed to an exceptional klutz, will become clear from the discussion in section 3.

[4] Zheng (2016), 62. Zheng's understanding is a helpful development of ideas initially articulated by Scanlon (1998) and Watson (2004).

[5] Zheng (2016), 63.

property but are unable to pay (or escape unidentified). This way of justifying strict liability can only be used for introducing a social practice for future cases. It would be unfair to retrospectively argue that such a rule would have been convenient, and we should therefore act as if it had been in place where the agents affected where not yet aware of such a rule and its demands on them.

This paper uses a second argumentative strategy. I rely on the idea that in some cases we can justify holding someone responsible for addressing a problem because their agency was the relevant cause of the problem and they are therefore the appropriate agent to hold responsible. More specifically, I discuss and adjust arguments holding that it is appropriate for others to hold us responsible for our actions not only in cases where the conditions of full *moral responsibility* apply but also in cases of what Tony Honoré and David Miller call *outcome responsibility*.[6] As will be explained in section 2, outcome responsibility refers to agency under conditions where the fallibility of the agents and/or the uncertainty of the conditions under which they act make it inappropriate to blame them for the accidents they cause but it nonetheless makes sense to attribute the outcome to their agency.

This discussion matters since cases of outcome responsibility describe scenarios very familiar to human beings. We are fallible beings acting under circumstances of uncertainty. Our capacities occasionally fail, and we act within a world of such complexity that it is unreasonable to expect us to ensure that none of our actions end up infringing the rights of other people.[7] Attributing duty-generating backward looking responsibility is important in cases like climate change where we face a problem and wonder who is responsible. With regard to historical emissions, for example, we can neither refer to pre-existing social conventions relating to strict liability nor build arguments based on culpable agency.[8] More generally, all too often fallible human beings acting under conditions of uncertainty cause considerable harm and the responsibility gap due to a lack of pre-existing social practices on the one hand, and demanding understandings of moral responsibility on the other, adds insult to injury for the victims. Outcome responsibility also matters for forward looking cases where we consider the normative dimensions of trying something untried like large scale geo-engineering to mitigate climate change.

Using Honoré's and Miller's ideas on outcome responsibility and Joseph Raz's idea of a sphere of secure competence, I argue that there is a meaningful way of considering people as responsible for outcomes caused by fallible beings in an uncertain world. The argument is that we should consider them as responsible in some cases of limited knowledge and capacity and liable *because* of this responsibility. The main reason why we should hold agents responsible in such cases is that the criteria for attributing agency and thus responsibility have to be realistic for human beings, who always have limited knowledge and whose capacities are not entirely reliable.

---

[6] The term *outcome responsibility* was coined by Honoré and further developed by Miller. See Honoré (1999), chapter 2 and Miller (2007), chapter 4.

[7] Cases where we fail to carry out actions the way we intend are even more numerous and harder to predict for cases of collective action.

[8] For accounts of why it is difficult to attribute moral responsibility for climate change, see e.g. Neumayer (2000), Gosseries (2004), Meyer and Roser (2010), Bell (2011), and Zellentin (2014).

The very idea of human agency assumes that fallible people in an uncertain world can be responsible – can be authors of change. Given the multiplicity of contributing factors involved in any human activity, attributing responsibility always involves an element of taking responsibility for a little more than what is fully under our control. As will be discussed, we have good reasons for attributing responsibility too early rather than too late.[9] In the following, I first explain my understanding of responsibility generally and outcome responsibility in particular (section 2). Section 3 discusses the agency of fallible agents. I start with presenting and criticising Honoré's and Miller's initial accounts and show how incorporating Raz's reference to a sphere of secure agency helps to better explain why we should – in some circumstances – consider agents to be responsible in a normatively significant way even though their actions are not blameworthy due to the failing of their normal capacities. Section 4 repeats this approach for the context of acting under uncertainty. The concluding section 5 emphasises that any plausible account of holding people liable based on their normatively significant backwards looking responsibility depends on an appropriate ethics of risk. The main normative point is to suggest that we ought to think ahead which conventions we might want to establish to balance incentives for innovation with a regime that adequately ensures that the victims of the unintended side effects of fallible agents acting in an uncertain world are not left to deal with the outfall unaided.

1 Understanding responsibility
Responsibility is a complex term in philosophical discourse. Kurt Bayertz introduces a helpful four-pronged relation to analyse the idea:
1. a *subject* (agent) is responsible
2. to an *addressee* (someone who is entitled to hold us responsible)
3. for an *object* of responsibility, due to
4. a set of evaluative criteria.[10]

(4) could be misunderstood since there are distinct questions requiring normative evaluation for each of these dimensions and thus four distinct sets of criteria:
1. Who can be a subject of responsibility of the relevant kind?
2. To whom are the relevant agents responsible?
3. What kinds of actions and behaviours can be the object of responsibility?
4. Why is the agent in question rightly considered the subject of responsibility for the object in question?

Given the context of his discussion, it is clear that Bayertz considers the last set of criteria most crucial for understanding the particular nature of responsibility, and this paper will focus on this set. It is important to stress that the criteria for attributing responsibility for

---

[9] The examples discussed in this paper are predominantly those discussed by Miller and Honoré referring to individual agency in the physical world. However, many of the cases where this kind of outcome responsibility matters also concern our agency in the *social* world. We often fail in what we set out to do because we misjudged our capacities to convince others to support or at least not hinder us. Furthermore, our agency often has unanticipated harmful effects on others not merely because we were ignorant about features of the physical world, but rather because we had an insufficient understanding of the social world and its complex interactions. By seeing fallibility and uncertainty as natural conditions of human agency, this paper sets out to show that even in cases where the conditions for blameworthiness do not apply, the agent is intimately linked to his acts and their consequences and thereby stands in a particular relationship to the victims of these actions. This relationship explains why it is appropriate for the victims to hold them responsible (to some degree).

[10] See Bayertz (1995), 15f. I will say nothing in this paper on the second dimension assuming that the main instance to whom we are responsible are the victims suffering the outcomes of our actions.

something to someone are *normative*. Attributions of responsibility (of any kind) are not ontological statements but elements of normative practice.[11] This means that we need a theory justifying under which conditions we *should* hold someone responsible for something and why. The relevant reasons, however, will be different depending on what kind of responsibility we are concerned with. Three kinds of responsibility focus on different kinds of objects:

1. Forward-looking responsibility assigns to an agent the task of taking care to ensure or maintain a particular state of affairs.[12]
2. Backward-looking responsibility attributes authorship of the outcome of an action.[13]
3. Remedial responsibility assigns duties to do something about a particular problem.[14]

The aim for this paper is to elucidate the notion of a rectificatory duty, understood as a version of remedial responsibility that is based on a particular kind of backward-looking responsibility. The claim is that in some circumstances an agent is remedially responsible for doing something about a problem – for rectifying a wrong – *because* they are responsible, in a normatively significant way, for bringing the wrong about. To put this into Bayertz's formula:

1. Agents are remedially responsible
2. to the victims of their action
3. for rectifying the problem they caused
4. if
   a. they really are responsible – in a normatively significant way – for bringing about this problem, and
   b. the problem causes the victims to suffer wrongful loss.[15]

The normative criteria informing (4) are taken from the usual understanding of rectificatory justice. This paper does not defend condition (4b). The idea that liability requires both responsibility and wrongfulness is deeply embedded in all debates on rectificatory justice and liability. The paper focuses exclusively on defending a particular understanding of condition (4a). Often, we assume that for agents to be responsible – in a normatively significant way – for bringing about a problem, they must be morally responsible for causing the harm. This requires the following:

---

[11] Following Bayertz, I will assume that this practice evolves and is at least party receptive to reasoning. For a detailed analysis of the development of different understandings of responsibility against the background of the social and political circumstances in which they emerged, see Bayertz (1995). For contemporary proposals as to when we should consider agents responsible, see e.g. Miller (2007)'s connection model, Young (2006, 2011)'s social connection model, and Goodhart (2017)'s political model of responsibility.

[12] Duff (1998) uses the term *prospective responsibility*, Raz (2011) calls this *responsibility₃*, and Birnbacher (2001) speaks of *ex-ante responsibility*.

[13] Duff (1998) here refers to *retrospective responsibility*, Raz (2011) to *responsibility₂*, and Birnbacher (2001) to *ex-post responsibility*.

[14] This usage relies on Miller (2007).

[15] Roughly, there are two ways in which a loss can be wrongful: a) if it is the result of a wrongful act, and b) if it infringes the rights of the victim despite being the result of a permissible act.

a) the agent could and should have known about the possibility of imposing the harm in question (knowledge condition),
b) the agent could have acted otherwise (capacity condition), and
c) it really was their action that brought about the harm (causality condition).[16]

The idea of outcome responsibility suggests that there are cases where normatively significant responsibility exists even though these conditions are not entirely fulfilled. To clarify, I want to distinguish between three kinds of backward-looking responsibility:

1. Causal responsibility: the agent played a central role in the causal chain of events that led to the outcome for which rectification is required. However, since this role does not refer to his agency, there is no normatively significant reason to attribute remedial responsibility based on this causal connection.
2. Outcome responsibility: the agent's agency played a central role in the causal chain of events that led to the outcome for which rectification is required, but while the agency involved provides sufficient reason to hold the agent remedially responsible, it is not the kind of agency that invokes blameworthiness.[17]
3. Moral responsibility: the agent's agency played a central role in the causal chain of events that led to the outcome for which rectification is required and is the kind of activity for which we blame to the agent.

In cases of outcome responsibility authorship of an undesirable outcome and blameworthiness come apart. This paper aims to disentangle the conditions for agency and blameworthiness in two scenarios:
1. Agents' capacities fail, and they accidentally wrong another person.
2. Agents act under circumstances where – despite their best efforts – they cannot know whether their precautions against unintended side effects will be sufficient to prevent wrongs to others and it turns out in retrospect that they did not.

In these cases we need to analyse why we should consider the conditions of genuine agency as fulfilled even though the conditions of blameworthiness are not.

2 The agency of fallible beings

Both Honoré and Miller discuss people with slightly restricted capacities, who innocently cause wrongful losses to others, as one example of agency without blameworthiness.

---

[16] The causality condition also raises a number of important questions and problems, which this paper does not engage with. Bayertz (1995) for example, points out that causality is not a fact of the world but an interpretation of the world, and that attributing moral significance to causality requires as much justification as any other proposed criterion for establishing a link between an agent and an outcome. Coleman (1992), chapter 14 raises doubts about the claim that causality is an obvious and necessary criterion for attributing the kind of responsibility that might relate to corrective duties. See also the problems raised in the discussions on causation in Cane (2001), Owen (2009) and Ripstein (2001).

[17] Unfortunately, in the course of Honoré's and Miller's writings, "outcome responsibility" refers sometimes to the particular kind of backward-looking responsibility with which we are concerned in this paper and sometimes to the particular kind of remedial responsibility that is grounded on this kind of outcome responsibility. This is a source of misunderstandings, since it implies that outcome responsibility is not only necessary but also sufficient for remedial responsibility, which is not the case. This paper is exclusively concerned with outcome responsibility as a kind of backward-looking responsibility.

Honoré speaks of "shortcomers".[18] Miller's example is an unusually clumsy person who cannot help causing havoc when entering a cut-glass shop.[19] This person is not at fault and should not be blamed, because their clumsiness is outside their control. The capacity condition is thus not entirely fulfilled while it is yet assumed that there is a sufficient degree of genuine agency to attribute normatively significant responsibility. In the following I will first present the original account, then highlight some problems with it, and finally show how these can be overcome with reference to Raz's work.

A. The original account
(i) Why is this a case of genuine agency?
By classifying the klutz in the cut glass shop as an example of outcome responsibility, Miller assumes that this is a case of genuine agency rather than mere inadvertent bodily movement.[20] Since the offending elbow seems to be a prime example of inadvertent bodily movement, this requires further explanation. Miller says: "We expect such a person to be aware of her shortcomings and therefore to stay away from shops full of cut glass."[21] That is, the focus shifts from the action of breaking the vase by inadvertently stretching out an elbow to the action of entering a cut glass shop even though one is fully aware of being clumsy.

(ii) Why does this kind of agency trigger responsibility?
Both Miller and Honoré argue that we should treat "shortcomers" as normal when it comes to attributing responsibility and thus assume a lack of care rather than incapacity as the origin of the accident in question.[22] Neither is terribly clear on this, but you can find in their explanations at least three distinct reasons for treating "shortcomers" as if they were normal:
1. Such a scheme is a requirement of consistency if we want to take credit for our good traits and their consequences. Our talents are, after all, as much a matter of luck as our handicaps.[23]
2. Such a scheme is overall beneficial for everyone.[24]
3. Such a scheme is more in line with our self-understanding as moral agents who are – despite all kinds of natural and social contingencies and pressures – still masters of our own lives.[25]

The argument from consistency does not require further explanation. The second line of reasoning is instrumental and emphasises that being considered responsible even for the

---

[18] Honoré (1999), 16.

[19] See Miller (2007), 88f.

[20] See Miller (2007), 87.

[21] Miller (2007), 89.

[22] Honoré (1999), 18.

[23] See Miller (2007), 87. Honoré makes this point the other way round. He argues that a scheme of strict liability can only be justified if it also ensures that we benefit from the consequences of our good traits. His approach is to defend the beneficial praxis of strict liability by arguing that such a scheme is justifiable if it fulfils the following three conditions: "the system must in its operation be 1. impartial, 2. reciprocal and 3. over a period, beneficial" Honoré (1999), 26.

[24] See Honoré (1999), 38 and Miller (2007), 89.

[25] See Honoré (1999), 29 and Miller (2007), 87.

consequences of one's clumsiness is beneficial both to klutzes and to their innocent victims. Being held responsible for the havoc one may happen to cause (when, for example, one enters a cut glass shop) is the price of being considered the kind of being who can legitimately choose to access spaces of a certain sort (e.g. such a shop). The alternative – that is, holding that any results of clumsiness should not be attributed to the agent – might give reason to limit what such (clumsy) people are free to do (unattended) for the sake of potential victims. A regime of attributing responsibility that treats them as if they were normal seems to balance their interest in freedom to enter shops of their choosing with the interest of shopkeepers in not suffering from other people's clumsiness.[26] For this paper, the third line of argument is the most interesting as it concerns the crucial link between agency and responsibility. The idea seems to be that – at least within the cultural context of modern Western democracies – being considered an independent and self-determined moral agent is so important to our self-understanding and self-respect as agents that we would rather take responsibility for something that it was not within our power to prevent than to admit that we are less than fully competent agents.

B. Open questions
(i) Negligence versus recklessness
A first problem emerges when we consider the following feature of the example: we assume that klutzes *can* compensate for the risk they pose to someone else's possessions when entering shops by paying for anything they happen to break. As mentioned, attributing responsibility here seems to balance their interest in freedom to enter shops with the interest of shopkeepers in not suffering from people's clumsiness. However, here the stakes are not very high. If the stakes are raised, things look different. Imagine a traffic accident where a driver had a seizure, lost control of his car, and killed pedestrians. Usually, we would consider the driver as a victim of their seizure and as not responsible at the time of the accident, given that they were not in control of their faculties at the critical time. If it turns out that the driver had been in accidents caused by such seizures before and thus should have known that they risked harming others, our intuition changes. In this case, we consider driving a car despite knowing that one's medical condition poses a threat to others as the relevant act and attribute not only outcome responsibility, but also blameworthy recklessness.

Miller explained the original example as a case of genuine agency by shifting the focus from agents' inadvertently stretching out their elbow to their choosing to insert themselves into a vulnerable environment. However, this move threatens to undermine the claim that this action is blameless. We often censure reckless behaviour; should we therefore consider this case one where blameworthiness comes into play, and thus one of moral responsibility? And if not, why not? The car accident case shows that we sometimes also have reason to treat "shortcomers" differently from normal people given that they pose unacceptable risks to others. The case is thus not a pure case of restricted or failing capacities but prompts considerations that will be discussed in more detail below in the context of acting under uncertainty – namely, the question what risks we can reasonably impose on others. We will postpone this discussion for now, but it is important to note

---

[26] Cane (2001), 106 claims that such a regime measuring everyone on a socially defined standard aims to strike a balance between our interests as agents and victims. Miller (2007), 87 promotes this as the fairest solution to resolve the competing interests of different agents.

that there are limits to when we can treat "shortcomers" as if they were normal. These limits have to be defended, and they have to be defended according to the criteria that determine how "normal people" are to be treated in cases of accidents.

(ii) What does it mean to treat "shortcomers" as if they were normal?
A second set of critiques suggests that Honoré and Miller do not say enough about the criteria determining how to treat accidents caused by "normal people". What is interesting about their discussion of responsibility in the case of restricted capacities is that they focus on the exception rather than the rule. "Shortcomers" in their sense are exceptions. By definition, most of us are capable of displaying "standards of care that a reasonable person of ordinary prudence can usually be expected to display".[27] Yet normal people too occasionally fail to display this standard of care. This is how accidents happen. Most of us slip up and break a glass occasionally, and when we do, we too are considered responsible but not blameworthy. We are not blamed for mishaps unless we were reckless. We are considered responsible even though we could not prevent it from happening. Try as we might, it is impossible for human beings to prevent slipping up every now and then. The responsibility attributed here thus seems to be a prime example of outcome responsibility. Why then focus on "shortcomers"? Miller needs to focus on "shortcomers" to make the move from focusing on the act in question (stretching out an elbow and knocking over a vase) to the deliberate prior act of entering a cut-glass shop in full knowledge that one is clumsy. For Honoré – as we shall see – it is far less clear why he explains outcome responsibility with regard to "shortcomers".

In any case, before arguing that we sometimes have good reason to hold "shortcomers" responsible in the same way as we do "normal people", we need an account of *how and why* we hold "normal people" responsible for accidents. If any kind of normatively significant responsibility requires genuine agency, then it is not clear at all why I should be responsible in cases where my usual capacities fail me. Honoré here refers to the distinction between one's general capacity and one's particular capacity in the given situation and argues that in order to be considered outcome responsible it suffices that one generally possess the minimal capacities necessary to be considered a responsible agent.[28] In justifying why this is fair, he claims that such a regime of responsibility benefits "normal people" more often than it fails them.[29] He thus makes the same instrumental argument for "normal people" as he does for "shortcomers", which raises the question: why focus on the latter?

(iii)   Why should the desire to be treated as if one were normal inform attributions of responsibility?
The most serious critique of both Honoré's and Miller's accounts of outcome responsibility in the context of "shortcomers" is that their strongest defences are ultimately instrumental. It is indeed beneficial to attribute responsibility when it is too costly or too invasive of privacy – in proportion to the damage in question – to determine whether someone was negligent, normally clumsy, or an exceptional klutz. However, this instrumental reasoning falls short in terms of Honoré's declared aim of establishing a

---

[27] This is how Coleman (1992), 218 puts it.

[28] See Honoré (1999), 11.

[29] See Honoré (1999), 26.

moral basis for strict liability.[30] The third kind of reason is different: "shortcomers" want to be considered like "normal" people and rather incur responsibility for acts outside their control than being considered agents with a limited capacity for responsibility. This would offer a moral reason intrinsic to the relationship between agency and responsibility. However, as Stephen Perry asks, why should the individual desire to be considered a fully competent agent (despite evidence to the contrary) justify a general social practice of attributing responsibility in cases of restricted capacity?[31]

C. The adjusted account

While these challenges mean that Honoré's and Miller's defences of outcome responsibility in the case of "shortcomers" leave some questions open, they point to some themes that are crucial for understanding attributions of responsibility in the context of restricted capacities:

1. In some cases agency based on deficient or failing capacities is still genuine agency.
2. The conditions for attributing blame are different from those for attributing agency.
3. The social practice of attributing responsibility for one's actions is closely linked to our understanding of competent agency and is thus important to an agent's self-understanding.

In the following, I will pick up on these themes to defend an adjusted account of outcome responsibility in the context of fallible agency. The most important move is to consider normal people whose capacities inexplicably fail rather than "shortcomers" as the prime example.

(i) Why is this a case of genuine agency?
We sometimes cannot help dropping a glass – even when we are careful and even when we are not particularly clumsy. Accidents happen. We know we should have been more careful because if we had been, the glass would not have dropped, but as often as not we would be hard pressed to explain just how we could have been more careful. Given that preventing such accidents is often out of our control, in what sense – if any – are such accidents still the result of our agency? Agency is often linked to control, and it is often thought that it is this control that links agency to responsibility. The control condition for responsibility suggests that I am responsible in a morally significant way only when it was in my control to bring about an outcome or to prevent it.

The point of this section is that for there to be genuine agency, and thus normatively significant responsibility, this control does not have to be perfect. Even when we conceptualise agency in a philosophical manner and lay out abstract criteria for what it means to deliberately act in the world, we must not idealise away the fact that human beings are fallible beings. The conditions to be applied in determining whether or not something is genuine agency or better described as "accidental bodily movement" must account for the fact that we are human beings whose capacities sometimes inexplicably fail. It is helpful to distinguish between three different ways in which what we intend to do can go wrong:

---

[30] See, for example, Cane (2001), e.g. 86 and 106.

[31] See Perry (2001), e.g. 63, 66.

1. We are in the midst of doing something that we know we are usually well able to accomplish, but our capacities fail.
2. We are trying to do something in full knowledge that we might well fail to accomplish our aim because we are unsure about the level of our relevant abilities to realise our aim in the circumstances at hand.
3. We are in the midst of doing something that we can usually accomplish without a second thought, but our usual capacities are suspended or interfered with.

In the third case our agency is undermined by external forces and therefore we do not incur responsibility. Apart from cases of fainting, etc., this also involves the responsibility-refuting conditions of derangement, manipulation, and coercion mentioned by Miller.[32] The second case will be analysed more closely in the next section on acting under uncertainty (section 4). Here, I am concerned with the first scenario.

When we do something like setting the table or walking, we know that we are doing something we are usually quite able to accomplish. Raz describes this as our "domain of secure competence". We are aware of "a domain within which we are confident that, barring competence-defeating events (a seizure, a biased teacher, etc.), if we set ourselves to do something we will".[33] However, it would be foolish for anyone to claim that, because they know how to set a table and how to walk, they will never accidently drop a glass or stumble. Raz dismisses accounts of responsibility based on control and instead proposes what he calls the "Rational Functioning Principle":

> Conduct for which we are (non-derivatively) responsible$_2$ is conduct that is the result of the functioning, successful or failed, of our powers of rational agency, provided those powers were not suspended in a way affecting the action.[34]

Raz thus distinguishes between cases where our capacities *fail* – and we are still responsible for the outcome of our actions – and cases where our capacities *are suspended* – and we are not. The key to distinguishing between these two cases lies in Raz's understanding of what it means to be a rational agent. According to him we are responsible in a normatively significant way when "our capacities of rational agency were available to us to guide and control our actions".[35] As rational agents, we know our domain of secure competence – that is, we know what we can and cannot do and achieve. However, the unspoken premise in Raz's discussion is that we know that we are human beings, not Gods, and that our capacities occasionally fail. That is something we have to keep in mind when contemplating what to do. This makes the case different from circumstances where our capacities are undermined or suspended due to extraordinary external circumstances that we cannot be expected to have considered.

(ii) Why is it inappropriate to attribute blameworthiness here?
When we blame someone, we think that the person whom we blame should have acted differently and that it reflects badly on their moral character that they did not. In the case of accidents that happen because our usually secure capacities fail, the agent's act does

---

[32] See Miller (2007), 91f.

[33] Raz (2011), 244f.

[34] Raz (2011), 231.

[35] Raz (2011), 246.

not reflect badly on their character, because while we think that they should not have caused the relevant losses, they could not have acted otherwise insofar as the critical element of the act in question was not under his control. The criteria for blameworthiness are more stringent than the criteria for agency.[36] Whereas the potential for failure with regard to our capacities is part of the very nature of human agency, the conditions for blameworthiness are not met in such a case. For something to reflect badly on our moral character, we must have had a genuine choice and opted to act in a morally problematic way.

(iii)   Why does this kind of agency trigger responsibility?
Raz argues that responsibility tracks cases were we act as rational agents – where we employ our capacities to understand and evaluate our situation, to guide our behaviour, and to execute the action we decide upon.[37] My interpretation of why this also includes cases where our capacities fail is that fallibility and vulnerability to accident are included in the nature of our capacities as human rational agents (and are included in our understanding of these capacities). When acting as rational agents, we act on this knowledge.

Raz, like Honoré and Miller, explains this with reference to the notion of "agent regret". This idea, first expressed by Bernard Williams, describes the fact that people identify with the consequences of their actions even when they know they were not at fault.[38] The authors imply that our social practice of attributing responsibility should take into account this identification because it is central to our understanding of our own agency. Discussions about "shortcomers" reveal this feature particularly clearly, since here the agents in question would seem to want society to disregard the fact that their capacities fail more often; they would thus seem to want to take responsibility for matters outside their control more often. "Shortcomers" want to be considered rational agents, and if this means taking on more responsibility for undesired consequences, so be it. As mentioned before, critics like Perry are reluctant to view this desire as a sufficient reason to justify a social practice of attributing genuine and morally significant responsibility (rather than merely providing instrumental reasons for a particular liability scheme). The point I wish to make is that the practice of attributing responsibility does not rest on the desire to be considered a rational agent, but on an understanding of agency that is informed by the knowledge that human beings are fallible beings. The discussions by Honoré, Miller, and Raz show that there are very good instrumental reasons to accept the costs associated with being considered a rational – responsible – agent. However, the argument for why agent regret should inform our practise of attributing responsibility does not rest on these reasons. Rather, it acknowledges that human agency is fallible and that any social practice based on agency needs to take this feature of human agency into account.

---

[36] Raz does not distinguish between morally significant responsibility involving blameworthiness (moral responsibility) and morally significant responsibility not involving blameworthiness (outcome responsibility). While his discussion is nuanced, he seems to be willing to bite the bullet and allow blaming those who fulfil the criteria of agency/responsibility even in cases of failed capacities. See Raz (2011), chapters 12 and 13.

[37] See Raz (2011), chapter 12.

[38] See Williams (1982), 123.

3 Acting Under Uncertainty

A second prominent example for outcome responsibility is agents taking a deliberate and reasonable gamble that they would be unlikely to cause wrongful losses to innocent third parties but ending up doing just that. Honoré speaks about "implicit bets".[39] Miller's example is one's bonfire accidentally setting the neighbour's shed on fire: "in lighting the bonfire I have taken a reasonable risk, but on this occasion I have gambled and lost."[40] As we shall see, the evaluation hinges on the epistemic conditions of agency.

A. The original account
(i) Why is this a case of genuine agency?

As before, Miller shifts the focus of the evaluation from the actual moment where the agent might well be unable to control the wind's blowing sparks across the fence to an earlier decision situation, where the agent decided to light the fire. If the conditions of agency applied in the situation of decision-making, then the agent deliberately made a bet, and it is this action that triggers responsibility for the unfortunate outcome.

(ii) Why is it inappropriate to attribute blameworthiness?

Miller's discussion of the example is brief, but he emphasises that the agent took a "reasonable risk". Which risk impositions are reasonable in any given contexts is a contested matter, but it is clear that Miller assumes that the agent in his example took all reasonable precautions to prevent the bonfire from spreading. Provided that the risk imposition indeed was reasonable (and thus does not reflect badly on the moral character of the agent) the agent remains blameless even when their bet goes wrong.

(iii)   Why does this kind of agency trigger responsibility?

Even when making a bet can be considered genuine agency, it is not obvious why doing so establishes a normatively significant relation between the agent and the consequences resulting from the bet. By its nature, a bet always involves an element of luck that is outside the control of the agent. The focus on agency in determining responsibility links back to the intuition that control over the situation matters when ascribing authorship. This intuition matters even if Raz is right to argue that the relevant understanding of the control links to rational agency rather than actual influence on the outcome. In situations where we make bets, we know that there is more than one possible result and that undesired consequences may obtain. We also know that we cannot ultimately control which outcomes actually obtain. To understand why Honoré and Miller think that the conditions of agency during the decision situation matter for ascribing authorship while the element of luck should be disregarded, it is necessary to look more closely at how they understand bets. Both emphasise that for something to be considered the making of a bet, and thus genuine agency, the agent must be able to evaluate the risks involved.

There is a normatively significant difference between cases where we know, for example, that lighting a bonfire might result in accidentally burning more than one intends to, and cases where the agent had no way of foreseeing that an action might have outcomes other than the one intended. A person waiting at the traffic lights wearing a vintage dress does not "bet" that she won't distract another driver due to her resemblance to his long lost love, thereby causing a horrible accident. Even should she somehow imagine such a

[39] See Honoré (1999), 25-27.

[40] Miller (2007), 93.

situation, it would still be inappropriate to describe her as making a bet – the role of coincidence is too great here to expect her to consider this as one of the possible side effects of her choice of dress. Both wearing an attention catching costume and going nude would be different.

Making a bet has epistemic preconditions. Not all instances of agency where we cannot be sure about the outcome qualify as bets and thus as deliberate action. As both Honoré and Miller emphasise, agency – even agency under uncertainty – has epistemic conditions. These conditions, however, do not primarily concern actual knowledge but rather "reasonable foresight". Bets apply to situations where agents

   a) can know what risks might be involved (if they bother to find out),
   b) can be expected to make themselves aware of what could go wrong, and
   c) ought to consider the possible consequences before deciding to act.

Given this normative and epistemic situation, agents can be seen as accepting the outcome of their bet by accepting the influence of luck when it comes to how their action might pan out. In their discussions of cases of "implicit bets" both Honoré and Miller thus argue that the fact that an agent was willing to make a bet establishes a morally significant link between them and the consequences of their bet (whether it goes well or not). They imply two distinct sets of criteria as significant for identifying this link:

   a) criteria concerning the choice situation where the agent (implicitly) makes the bet,[41] and
   b) criteria to determine whether the act in question was a genuine bet.

The latter set of criteria becomes more important the more implicit the bet is. As we have seen, these criteria are partly epistemic: it has to be possible for the agent to know and evaluate the relevant outcomes. However, these criteria have a significant normative dimension: what matters is not what the agent actually knows about his choice situation. Rather, what matters is what they should know.[42]

B. Open questions
(i) Why link agency to reasonable foreseeability rather than actual foresight?
However, Arthur Ripstein argues that what matters when it comes to *attributing* agency must be actual knowledge rather than what the agent *should* know. He claims: "Something cannot be an expression of *my* agency if it played no part in my practical deliberation, even if others might have thought of it."[43] If Ripstein is right, then Honoré's and Miller's accounts of outcome responsibility of the "implicit bets" kind are unduly extensive and include cases where the relevant kind of agency is lacking.[44]

---

[41] This primarily refers to the familiar requirement that there be no derangement, manipulation or coercion.

[42] See Miller (2007), 96f.

[43] Ripstein (2001), 103.

[44] This does not rule out blameworthiness, because we can blame agents for disregarding important considerations that they were duty-bound to consider. It is important to note, however, that the blameworthiness here does not refer to the contribution to the harm that came to pass but to the failure to abide by duties of care.

(ii) Are we really only responsible for unintended consequences in cases where we know precisely what risks are involved in what we do?

Another critique, however, might state the exact opposite and argue that Honoré's and Miller's accounts of responsibility in the context of implicit bets are unduly narrow and exclude important cases where we might well consider choices under uncertainty to generate normatively significant responsibility for unwanted outcomes. Honoré emphasises at several points that where we attribute outcome responsibility (and later liability based on this responsibility), the harm that came to pass has to be of the kind that was considered in the initial risk assessment.[45] On a strict reading this means that we can only be responsible in a normatively significant way for precisely those bad consequences that we could and should have anticipated. We can be responsible for unintentionally lighting the neighbour's shed because losing control of fire is a commonly known risk. We cannot be held responsible for historical greenhouse gas emissions, because until fairly recently we could not have known that such emissions might accumulate in the atmosphere and affect the global climate. This seems convincing until we look at the implications of this interpretation for acting under uncertainty more generally. The question is whether ignorance of the possible risks should absolve us in all instances of agency under circumstances of uncertainty. Below I will argue that where we know that we do not know what risks might be expected, and know that we cannot rule out the possibility that our actions will put others at risk, special duties of care exist.

C. Agency in the Context of Uncertainty

In the following I extend Honoré's and Miller's ideas on implicit bets to explain responsibility in the context of agency under conditions of uncertainty more generally while at the same time avoiding Ripstein's challenge with reference to Raz's work.

(i) Why is this a case of genuine agency?

My case against Ripstein's challenge is that the epistemic condition for genuine agency is fulfilled as soon as the agent knows whether something falls within or outside of her sphere of secure competence. Given Raz's definition of this sphere, this must be something of which agents are aware whenever they are acting rather than accidentally moving body parts or suffering from delusion or manipulation. As mentioned, Raz identifies the sphere of secure competence as the domain of acts that we know ourselves to be able to perform successfully, barring either the failure of or interference with our capacities. In the last section, I argued that when we act as rational beings we know that, as human beings, we are fallible and need to take this into consideration. In this section, I add that as rational beings we know that when we act in the world nothing is ever one hundred percent certain. The knowledge that we usually succeed in doing what we set out to do when acting within our domain of secure competence is thus unlikely to attach only to acts where absolutely nothing can go wrong unless something drastic were to happen. In the real world, there are very few such acts. Rather, I propose that it should attach to those acts where we know the associated risks so well that we do not have to re-evaluate them in each instance, and also know without further thought how to limit the remaining risks in such a way as to make their imposition on others reasonable and legitimate.

---

[45] See e.g. Honoré (1999), 27.

My claim is that "lighting bonfires" might be something that is in a person's sphere of secure competence even though freak accidents due to stray sparks happen. The person in question would have to have the experience with lighting fires that qualifies them to be sure that they have anticipated all common risks associated with fire and have taken all reasonable precautions. Whenever a person acts in the knowledge that what they do falls within their sphere of secure competence, they make the implicit bet that what is usually enough in terms of safeguarding against unwanted side effects is actually enough. The person might not consciously deliberate on the details, but given that they know what they are doing in the minimal sense of knowing that their action falls within their sphere of secure competence, their action in these circumstances qualifies as agency.

Far more interesting, and far more common, are cases where we act in full knowledge that we are acting outside of our sphere of secure competence. For the majority of us, lighting bonfires in a manner that involves taking all reasonable precautions requires conscious effort. We need to research and evaluate the risks of burning this kind of material in this kind of surrounding and weather, etc. And while most people know enough about fires to ensure the reasonable safety of their bonfires, they know that fires never are perfectly safe and that if they choose to light one, they deliberately bet that their precautions will suffice. This bet might be implicit; the agent might not explicitly construct a line of thought along the lines of "Well, I did what I could and I bet that's going to be enough". It is nonetheless a deliberate bet given that the immediate knowledge that this act is outside one's sphere of secure competence coincides with the knowledge that risks to others might exist.

(ii) Why is it inappropriate to attribute blameworthiness?
If we choose *not* to conduct a risk assessment, we invite charges of recklessness and thus blameworthiness. When we do conduct a proper risk assessment and come to the conclusion that as far as we know the risks we impose are reasonable and legitimate, we avoid blemishing our moral character even if things go wrong.

(iii)   Why does this kind of agency trigger morally significant responsibility?
We do not, however, escape the normatively significant link between our action and its implications that is established by the choice to impose these risks on others. This decision might be implicit, but it is nonetheless deliberate whenever we know either that we are acting within our sphere of secure competence (which involves knowing the relevant risks) or outside of this sphere (which involves knowing that there are some things we do not know and control). The general idea is that genuine agency is also possible under conditions of uncertainty where we do not know the exact risks associated with our options. The relevant condition for agency – apart from the absence of coercion, derangement, and manipulation – is the minimal epistemic condition that we know whether or not what we are attempting to do is something we know we can do.

Genuine agency unproblematically establishes a normatively significant link between the agent and the consequences of her agency where the agent knows what kind of consequences she might bring about and is thus under a duty of care to evaluate the normative implications of the risks involved. Not knowing can also have normative significance. This is so in cases where we do not even know what it takes to conduct a reasonable risk assessment. Currently, we know that we do not know enough about the possible consequences of geo-engineering to be reasonably sure that we have taken all reasonable precautions once we engage in it. The argument here is that this knowledge

about the limits of our knowledge means that where we engage in geo-engineering, we cannot use our ignorance about the possible risks to absolve us from responsibility for the results or from the normative significance of acting in such circumstances. When we try something we know we do not understand, we take a deliberate gamble with regard to the well-being of others.

Freak accidents are relevantly different. We do not place a bet when we assume that a freak accident is not going to happen, because we can be reasonably sure that it is very unlikely that such an accident will happen. When we talk about freak accidents – like a vintage dress causing a car accident – we have a good enough understanding of the situation to categorise some outcomes as freakish. Where we know, however, that our understanding of the situation is limited, we do take a gamble; here the outcomes are not freakish – they are simply unpredicted. That is, outcome responsibility is different from moral responsibility to the degree that it does not require that the agent acted in a blameworthy manner. At the same time, outcome responsibility is different from mere causal responsibility since it disregards cases where the agents' contribution cannot be considered an expression of their genuine agency.

4   Conclusion

This paper argued that the conditions for attributing morally significant responsibility and blameworthiness can come apart in cases where fallible agents acting in an uncertain world cause undesired outcomes. It assumes further that under additional conditions – most importantly the fact that these outcomes infringe the rights of innocent others – this responsibility could be the basis for rectificatory duties even where there was no social practice establishing such liability beforehand. Most importantly, it suggests that introducing and continuing social practices that impose liabilities on agents who accidentally cause harm to others is appropriate.

How such practices should be designed in detail raises further questions. In many cases it might be inappropriate to consider the innocent harm bringers as the only agents responsible for restitution. In the nightmarish scenario where geo-engineering would be our only chance to prevent dangerous levels of climate change, we would have very good reasons to try something untried for the potential benefit of many. Attributing all the liability for undesired side effects to the agent making such a (potentially tragic) choice would hardly seem fair. Similarly, while the above theory allows for seeing the decision to allow air pollution to increase exponentially during the Industrial Revolution as a deliberate gamble and thus as outcome responsibility incurring agency, this does not mean that it is appropriate to distribute all climate change related duties based on the contributor pays principle. [46]

Nor would basing all liability on outcome responsibility make for a sustainable social practice as it could render attempts to innovation prohibitively expensive. When looking forward and consider the implications of future innovation it is thus reasonable to think about insurance schemes as well as public safe-guards balancing the public interest in innovation with the unknown risks associated with it.

---

[46] For more detailed discussions, see the literature discussed in footnote 7.

Furthermore, no liability scheme would be practically workable nor could it ever be fair unless it engages with an ethics of risk to decide which risks are deemed acceptable. When discussing how to treat accidents that are the result of people's normal capacities to fail (e.g. a minimal lapse in attention resulting in a car accident), social practices need to find a threshold to distinguish acceptable from unacceptable risks. Sometimes people whose capacities fail more often have to be treated differently because they pose unacceptable risks to others. Most people, for example, are susceptible to instances of microsleep and thus pose a certain risk to others when driving, but the risks associated with drivers suffering from epilepsy are considerably larger and thus justify certain restrictions. Evaluating risk impositions and identifying thresholds is extremely difficult as all major moral theories struggle with the moral assessment of acts that might or might not pose a danger to others.[47]

For the context of agency under uncertainty an ethics of risk is even more crucial and even more difficult. Evaluating unknowable risks is altogether impossible. If we really cannot know what dangers might be the undesirable side effects of an action and/or how unlikely undesired side-effects might be, we have no way to evaluate these risks impositions before they penned out. Any ethics of risks must demand that we make reasonable efforts to find out which risks we are imposing. This – sometimes – allows identifying unacceptable risks but sometimes we really cannot know. The recommendation for agency under this kind of uncertainty is that we make reasonable efforts to keep investigating potential risks as we carefully take step by step. Had we heeded the emerging indications regarding the potential threats associated with the greenhouse effect early, investigated thoroughly as soon as they emerged, and in response to the raising evidence aimed to move away from a carbon intensive industry decades ago, we would likely face fewer problems now.[48] With regard to heavy impact actions like some forms of geo-engineering where one shot at it might be enough to change everything quite drastically and for a very long time, it seems reasonable to demand that we need very good reasons for going ahead. There might be circumstances, where we have very good reasons to do so. But it seems reasonable and fair to demand that when we do, we have taken precautions and set resources aside to make sure that we are in a position to help those whose rights might be infringed. It seems also reasonable to suggest that those who make this decision, take some responsibility towards those who did not get to have a say. However, at the same time, from a normative perspective as well as from the point of view of justifying our social practices, it seems desirable that where decisions are made that are likely to impact on a lot of people, a lot of people are involved in the decision making process.[49]

This paper offers a way to justify social practices of attributing responsibility and liability for fallible agents acting in an uncertain world. It concludes claiming that the particular

---

[47] See Fleming, et al. (no year) for a brief but on the point summary of the relevant debates.

[48] This does not only, nor necessarily, address the CO2 accumulation in the atmosphere but also the degree to which our economy and social structures depend on carbon intensive technologies and infrastructure – the longer a system was run on a particular energy source and thus optimised life around it, the harder it is to change...

[49] See e.g. Fischhoff (1994) for a detailed discussion of the intersection between the ethics of risk and demands for democratic decision-making and Oberdiek (2012) for an account that sees the wrong of taking the risk of injuring another as a limitation of that person's autonomy.

form of these social practices always has to bear in mind an ethics of risk and thus justify which risks we want to consider as acceptable and on which criteria.

Bibliography:

Bayertz, K. (1995) 'Grenzen der Verantwortung'. In: Bayertz, K. (ed.) *Verantwortung: Prinzip oder Problem?* Darmstadt: Wissenschaftliche Buchgesellschaft.

Bell, D. (2011) 'Global Climate Justice, Historical Emissions, and Excusable Ignorance'. *Monist*, 94, 391-411.

Birnbacher, D. (2001) 'Philosophical Foundations of Responsibility'. In: Auhagen, A. E. and Bierhoff, H.-W. (eds.) *Responsibility. The Many Faces of a Social Phenomenon*. London / New York: Routledge.

Cane, P. (2001) 'Responsibility and Fault: A Relational and Functional Approach to Responsibility'. In: Cane, P. and Gardner, J. (eds.) *Relating to Responsibility*. Oxford: Hart Publishing.

Coleman, J. L. (1992) *Risks and Wrongs*, Cambridge: Cambridge University Press.

Duff, R. A. (1998) 'Responsibility'. In: Craig, E. (ed.) *Routledge Encyclopedia of Philosophy Vol. 8*. London: Routledge.

Fischhoff, B. (1994) 'Acceptable Risk: a Conceptual Proposal'. *Risk: Health, Safety & Environment*, 5, 1.

Fleming, J. I., Pike, G. K. and Ewing, S. (no year) 'The Ethics of Risk', part of: *Bioethics Resources* (Southern Cross Bioethics Institute), available at: http://www.bioethics.org.au/Resources/Online%20Articles/Other%20Articles/The%20ethics%20of%20risk.pdf [Accessed 17. July 2017].

Goodhart, M. (2017), 'Interpreting Responsibility Politically', *Journal of Political Philosophy* [online], doi: 10.1111/jopp.12114.

Gosseries, A. (2004) 'Historical Emissions and Free-Riding'. *Ethical Perspectives*, 11, 36-60.

Honoré, T. (1999) *Responsibility and Fault*, Oxford / Portland, OR: Hart Publishing.

Meyer, L. H. and Roser, D. (2010) 'Climate Justice and Historical Emissions'. *Critical Review of International Social and Political Philosophy*, 13, 229-253.

Miller, D. (2007) *National Responsibility and Global Justice*, New York: Oxford University Press, USA.

Neumayer, E. (2000) 'In Defence of Historical Accountability for Greenhouse Gas Emissions'. *Ecological Economics*, 33, 185-192.

Oberdiek, J. (2012) 'The Moral Significance of Risking'. *Legal Theory*, 18, 339-356.

Owen, D. G. (2009) 'Figuring Foreseeability'. *Wake Forest Law Review*, 44.

Perry, S. R. (2001) 'Honoré on Responsibility for Outcomes'. In: Cane, P. and Gardner, J. (eds.) *Relating to Responsibility*. Oxford: Hart Publishing.

Raz, J. (2011) *From Normativity to Responsibility*, Oxford: Oxford University Press.

Ripstein, A. (2001) *Equality, Responsibility, and the Law*, Cambridge: Cambridge University Press.

Scanlon, T. M. (1998) *What We Owe to Each Other*, Cambridge, MA: Harvard University Press.

Watson, G. (2004) *Agency and Answerability: Selected Essays*, Oxford: Clarendon Press.

Williams, B. (1982) *Moral Luck*, Cambridge: Camebridge University Press.

Young, I. M. (2006) 'Responsibility and Global Justice: A Social Connection Model'. *Social Philosophy and Policy*, 23, 102-130.

Young, I. M. (2011) *Responsibility for Justice*, Oxford / New York: Oxford University Press.

Zellentin, A. (2014) 'Compensation for Historical Emissions and Excusable Ignorance'. *Journal of Applied Philosophy*, 32, 258–274.

Zheng, R. (2016) 'Attributablity, Accountability, and Implicit Attitudes'. In: Brownstein, M. and Saul, J. (eds.) *Implicit Bias and Philosophy, Volume 2*. Oxford: Oxford University Press.