



On the Harm of Imposing Risk of Harm

Kritika Maheshwari¹ 

Accepted: 24 August 2021 / Published online: 24 September 2021
© The Author(s) 2021

Abstract

What is wrong with imposing pure risks, that is, risks that don't materialize into harm? According to a popular response, imposing pure risks is *pro tanto* wrong, when and because risk itself is harmful. Call this the Harm View. Defenders of this view make one of the following two claims. On the Constitutive Claim, pure risk imposition is *pro tanto* wrong when and because risk constitutes diminishing one's well-being viz. preference-frustration or setting-back their legitimate interest in autonomy. On the Contingent Claim, pure risk imposition is *pro tanto* wrong when and because risk has harmful consequences for the risk-bearers, such as psychological distress. This paper argues that the Harm View is plausible only on the Contingent Claim, but fails on the Constitutive Claim. In discussing the latter, I argue that both the preference and autonomy account fail to show that risk itself is constitutively harmful and thereby wrong. In discussing the former, I argue that risk itself is contingently harmful and thereby wrong but only in a narrow range of cases. I conclude that while the Harm View can sometimes explain the wrong of imposing risk when (and because) risk itself is contingently harmful, it is unsuccessful as a general, exhaustive account of what makes pure imposition wrong.

Keywords Risk · Harm View · Wrong-makers · Autonomy · Preference

1 Introduction

Philosophers discussing the ethics of risk imposition typically focus on cases of 'pure risks' where the risked harm fails to materialize (Thomson 1986).¹ Here are two familiar examples. Bill drives recklessly, risking the lives of pedestrians, but luckily nothing happens. Bill plays a game of Russian Roulette with his friend, but luckily the friend survives.

¹I only focus on cases of risk of *harm*. Moreover, for the purposes of this paper, I use the term 'pure risk' interchangeably with 'imposing pure risk', 'pure risk imposition', and 'risking harm'.

✉ Kritika Maheshwari
k.maheshwari@rug.nl

¹ Department of Ethics, Social and Political Philosophy, Faculty of Philosophy, University of Groningen 52, Oude Boteringestraat, 9712 GL Groningen, Netherlands

In such cases, Bill could have actually hit one of the pedestrians, or killed his friend. But Bill did none of these things. Yet, what Bill does seems wrong. What, then, explains it? According to one popular response, Bill's pure risk imposition is *pro tanto* wrong when and because risk itself is harmful, regardless of whether the risked harm materializes.² Call this the Harm View. Defenders of this view can be interpreted as making one of the following claims³:

On what I call the Constitutive Claim, pure risk imposition is *pro tanto* wrong when and because risk *constitutes* or counts as making the risk-bearer intrinsically worse-off in some respect. In arguing for this claim, influential accounts locate the constitutive harm of risk in risk-bearer's well-being being diminished viz. preference-frustration (Finkelstein 2003), or in their legitimate interest in autonomy being set-back (Oberdiek 2009, 2012, 2017).⁴ By contrast, on what I call the Contingent Claim, pure risk imposition is *pro tanto* wrong when and because risk has harmful consequences for the risk-bearer, for instance, psychological distress, such as fear and anxiety.

The Harm View offers a sufficient (though not necessary) condition for explaining the wrongness of imposing pure risks on others. That is, if risk itself is constitutively or contingently harmful, then this generates a moral reason against Bill's risk imposition that can be outweighed by other countervailing considerations. While some philosophers find the view appealing, others have expressed doubts.

For instance, some find it puzzling how imposing pure *risk* of harm on others can count as the same thing as harming them (Perry 2007; Gardner 2017, p.77). As Richard Arneson (2010) puts it, "[T]he sheer risk of suffering harm is just that, a risk of suffering harm, and not itself a harm." (p.346). If it is, then it is unclear how the harmfulness of risk itself should be discounted from the well-being of those for whom the risk never materialized. Some question whether this is even possible without running into circularity or without double-counting harm (Hurd 1996; Finkelstein 2003).⁵

This paper argues that the Harm View is plausible only on the Contingent Claim, but fails on the Constitutive Claim. In discussing the Constitutive Claim, I argue that both the preference and autonomy account fail to show that risk itself is constitutively harmful and thereby wrong. In discussing the Contingent Claim, I argue that risk itself is contingently harmful and thereby wrong but only in a narrow range of cases. I conclude that while the Harm View can sometimes explain the wrong of imposing risk when (and because) risk itself is contingently harmful, it is unsuccessful as a general, exhaustive account of what makes pure imposition wrong.

The paper proceeds in five sections. Section 1 covers two important preliminaries. I offer clarification of key concepts and then discuss the significance of Harm View in our practical deliberation. Section 2 tackles the Constitutive Claim by focussing on the preference account as developed by Claire Finkelstein (2003). Section 3 tackles the Constitutive Claim by focussing on the autonomy account as developed by John Oberdiek (2009, 2012). Section 4 discusses the Contingent Claim. Section 5 concludes.

² Henceforth, I will occasionally drop the term '*pro tanto*' for simplicity.

³ This distinction is often overlooked in recent discussions of the Harm View. C.f. Oberdiek (2017) and Parr & Slavny (2019).

⁴ For the purposes of this paper, I only focus on the preference and autonomy account. Other closely-related accounts that argue for the Constitutive Claim include Adriana Placani's (2016) dignity account, Jonathan Herington's (2019) and Seth Lazar's (2019) security account, and Maria Ferretti's (2016) freedom account. Wherever necessary, I've indicated where my discussion applies to these alternative accounts too.

⁵ For responses to these last two objections, C.f. Placani (2016). Others who also doubt the plausibility of the Harm View for different reasons include Judith Thomson (1986), Stephen Perry (2007, 2014), Joseph Bowen (unpublished ms.).

2 Preliminaries

Before we proceed, clarifying key concepts is essential. In particular, we need a conception of *harm* and *risk*.

First, consider harm. Defenders of the Harm View hold that the harm of risk is both conceptually and ontologically distinct from the harm that is risked. To distinguish between the two, theorists often employ the terminology of ‘risk-harm’ to refer to the former, and ‘outcome-harm’ to refer to the latter (c.f. Finkelstein 2003; Placani 2016). As the central focus of the Harm View is risk-harm (constitutive or contingent), theorists must specify its two components: the currency and the measure of harm (Tadros 2016, p.175–200). The former concerns the subset of bad things that can happen to the risk-bearer. The latter concerns the degree to which someone is harmed. As we’ll see, the precise contours of these two components are drawn differently on different accounts.

Second, consider risk. On the standard account, risk simply is the conditional probability of a bad outcome for someone. Probabilities come in two broad varieties, namely, objective and epistemic probabilities (Gillies 2000). Objective probabilities are determined by physical features of the world. They are mind-independent and typically interpreted in terms of relative frequencies. By contrast, epistemic probabilities are mind-dependent and denote either an agents’ degree of belief or degree of support given by available evidence for a hypothesis (Roser 2017). In arguing for the Harm View, most theorists take the salient probabilities to be objective ones.⁶ For now, let’s assume that in the cases that interest us, the relevant objective probabilities are available.

If it is wrong to impose risks on others in virtue of risk-harm, this carries both theoretical and practical implications. First, the notion of harm is central to normative, political, and legal theorizing (Hanser 2013). For instance, distinct versions of the so-called harm principle are aimed at criminalizing harmful conduct. According to John Stuart Mill’s (1859) canonical formulation, the only purpose for which power can be rightfully exercised over an individual, against his will, is to prevent harm to others. Ordinarily, Mill’s harm principle applies when one suffers actual or expected outcome-harm. If the Harm View is correct, there is a *prima facie* case for the principle to warrant restricting or constraining risk impositions in light of actual or expected risk-harm, in addition to the actual or expected outcome-harm one might suffer.

Second, we routinely confront and manage risks of falling ill and infecting others, or the risk of losing our job during a financial crisis, or the risk of facing life-threatening accidents in an environmental disaster, and so on. If the Harm View is correct, then governments, health organizations, insurances companies need to take account of risk-harm, in addition to the actual or expected outcome-harm that individuals may suffer when making decisions about safety policies, distributing benefits and risk, and so on.

Some have even suggested that risk-harm is crucial for thinking about compensating individuals for living with various climate-change risks (Herington 2017). Although climate change exposes many current and future generations to various risks of climate-related losses, a relatively smaller number would actually incur any real, outcome-harm. However, if the Harm View is correct, then compensation for harm should not be limited to those few who suffer outcome-harm, but should also be extended to the large majority of those individuals

⁶ For a critique of the Harm View for its failure to account for epistemic probabilities, c.f. Parr and Slavny (2019).

who suffer risk-harm but are lucky enough not to suffer outcome-harm, all other things being equal.

Third, besides questions of compensation, the plausibility of the Harm View is also relevant for questions of liability and punishment (Perry 1997; Finkelstein 2003). Within both tort and criminal law, wrongful harming, rather than harmless wrongdoing, is a necessary condition for judgements of legal responsibility for exposing others to risk. If the Harm View is correct, then risk-harm itself could serve as a basis for holding individuals liable as well as imposing sanctions for wrongful behavior in well-known cases of ‘victimless crimes’ such as dangerous driving, careless handling of firearms, and so on.

These are some important reasons that motivate a thorough discussion of the Harm View. Let’s begin with examining the Constitutive Claim.

3 The Constitutive Claim: Risk’s Impact on Well-Being

The Constitutive Claim holds that pure risk imposition is *pro tanto* wrong when and because risk *constitutes* or counts as making the risk-bearer intrinsically worse-off in some respect. What exactly is it to make someone worse-off? One response is that it involves diminishing their well-being. We can spell this out further in terms of a substantive theory of well-being. According to Claire Finkelstein (2003), a plausible candidate is the preference satisfaction theory.⁷

On a standard ‘actualist’ version of this theory, one’s well-being at a given time does not merely depend on one’s experiential state such as psychological distress or the achievement of certain objectively valuable goods. Instead, any increase or diminishment of an agent’s well-being is determined by the extent to which the world or state of affairs at a given time is or becomes how she *actually* prefers it to be (Heathwood 2006). By focusing on the risk-bearer’s actual preferences, Finkelstein (2003) argues that we can capture the wrongful harm of imposing pure risk in “a perfectly commonsensical way” (p.974).

To illustrate, she offers the following example (p.967). Suppose Bill fails to repair his airplane’s engine and thereby flies a defective airplane. His negligence exposes each passenger to a heightened risk of dying in a plane crash. At the moment of risk imposition, say at time t , the passengers belong to a reference class of people in which the relative frequency of dying in a defective airplane is greater than the reference class of people who fly in non-defective airplanes. According to the preference account, exposing passengers to the risk of dying *entails* a reduction in their well-being (that is, harms them) at t because they prefer to be in the latter reference class (facing lower or negligible risk of dying) rather than the former (facing higher risk of dying). Thus, risk *itself* is constitutively harmful by virtue of frustrating one’s preferences, thereby rendering Bill’s imposition of risk wrong.

Finkelstein further notes that on an *ex post* judgement of the case, say at some subsequent time $t + 1$, the harm of risk persists or “does not evaporate the moment it is clear no outcome

⁷ Finkelstein (2003) starts out by spelling harm in the currency of legitimate interest, but later states that she uses the term ‘interest’ as a synonym for ‘preference’ (p. 972). She ultimately connects this view with an actualist preference satisfaction view of well-being by noting that “an agent has an interest in all and only those items that correspond to something he wants or desires” (p.972). For this reason, I subsume her view under well-being account instead of the interest-based account that I discuss in the next section.

harm will result” (p.967). However, whether this is really the case depends on whether the passengers really suffer a genuine frustration of some preference *ex ante*, that is, at *t*.⁸ Inquiring this requires us to first identify the frustrated preference in question.

As Finkelstein relies on an actualist preferentialist theory, this preference must be intrinsic rather than instrumental.⁹ On this theory, only preferences for things that are for their own sake count towards one’s well-being at *t* (Parfit 1984; Heathwood 2006). Moreover, the theory holds that one is harmed if it is both true that one prefers that some state of affairs (that is, the object of one’s desire) at *t* obtain, and this state of affairs fails to obtain concurrently at *t* (Heathwood 2006; Sarch 2013).

Keeping these aspects in mind, there are at least two preferences that qualify as plausible candidates for what makes risk constitutively harmful: (i) one’s general, intrinsic preference at *t* not to be exposed to risk *per se*; or (ii) one’s general, intrinsic preference at *t* not to suffer outcome harm at *t* or at any subsequent time, say at *t* + 1. Below, I argue that for both these possibilities, the preference account fails to show that risk itself is constitutively harmful, thereby undermining its plausibility as an account of explicating the wrong of imposing risks in terms of risk-harm.¹⁰

Let’s begin with (i). It is plausible that many individuals possess a preference to be free from risk at any given moment in time. For instance, the fact that we buy health and travel insurances, wear safety belts, build earthquake prone homes, and so on, is indicative of that. Acknowledging this, some theorists think that the preference account rightly captures the harm of being exposed to risk at *t* (Herington 2019).¹¹ However, other things being equal, such an account implausibly entails that individuals are continuously made worse-off (harmed), *given* the widely held assumption that virtually every action we perform carries some risk for others. We quickly run into a kind of paralysis¹²: if it is simply impossible to ensure non-frustration of individuals’ preference to be free from risk at any given moment in time, we cannot help but constantly harm others by performing our most mundane, everyday actions.

At first glance, this merely implies that we have a *pro tanto* reason against most of our everyday actions, insofar in performing these actions we frustrate other’s preferences to be free of risk *per se*. If the preference account has this implication, it doesn’t necessarily make the account implausible.¹³ The underlying issue here, however, is that the preference account renders most of ordinary actions as instances of harm and thereby wrong. Even those who are willing to accept that our ordinary actions are *pro tanto* wrong by virtue of carrying a small or

⁸ In contrast, Herington (2017) thinks Finkelstein is making a claim about *ex post* preference to have been free from risk.

⁹ In the discussion that follows, I only focus on actual preference account as that is the focus of Finkelstein’s account. Others may appeal to some version of an ideal or fully informed preference theory. On this theory, the relevant preferences are ones that an idealized agent would possess were she to be fully rational and had full information about the risks. My objections to the preference account hold even if one were to appeal to ideal rather than actual preferences about risks.

¹⁰ Others have also objected to Finkelstein’s view for different reasons. For instance, John Oberdiek (2012) objects that the harm identified by preference frustration is of the wrong kind. Jonathan Herington (2019) objects that Finkelstein’s account faces a dilemma between embracing an implausible actual-desire theory or embracing an ideal theory that fails to account for risk.

¹¹ Herington (2019) thinks this is plausible on a general description of the preference theory.

¹² Hayenhjelm and Wolff (2011) raise a similar problem of paralysis for rights-based theories. The idea there is that it is impossible to assign individuals the right not to be subjected to risk without their consent, insofar virtually every action carries with it some risk, however small, of serious harm to others.

¹³ I thank an anonymous reviewer for pressing me on this point.

even a serious risk of harm for others might find it unacceptable that an account renders our ordinary actions wrongful *because* harmful for at least two related reasons.

First, it seems to stretch the notion of harm beyond its ordinary limits. It treats too wide a range of cases as instances of harming. In doing so, it absolves the relative seriousness of imposing outcome-harm on someone, much against our commonly held intuitions. For instance, causing someone excruciating pain is much more serious than causing them to suffer risk-harm by virtue of frustrating their preference not to be exposed to an increased risk of suffering excruciating pain.

And second, the account risks undermining or trivialising the special moral significance many theorists attribute to harm (Shiffrin 2012). Most agree that we have strong reasons, and often, even a strong obligation not to harm others. However, the preference account problematically entails that in performing ordinary actions, we fail to observe this duty by virtue of harming others *viz.* frustrating their preference not to be subject to risk *per se*.

Perhaps we can simply modify (i) and concede one's preference against *specific* risk under *specific* conditions as the relevant candidate. For instance, suppose we only care about a passenger's preference not to be exposed to *non-consensual, higher or increased* risk of dying in a defective airplane crash. This narrows the range of relevant cases without leading to paralysis as before. Still, a number of problems remain.

First, the account makes the constitutive harm of risk *contingent* on agents' having preferences about the specific risks they face. This makes the account extensionally inadequate in at least two kinds of cases: (1) cases where risk-bearers have no actual preferences about the risk in question; and (2) cases where we simply cannot appeal to risk-bearer's preferences. As an illustration of the first case, think of a newly born baby who is on board and also exposed to the risk of dying in a plane crash because of Bill's conduct. The baby, however, does not herself have any actual preferences for or against being subject to risk. As an illustration of the second case, think of currently non-existing future generations who are subject to climate change risks because of our actions. We do not know anything about their preferences regarding these risks (and also in general).¹⁴ In these cases, there is a strong intuition that imposing risk is wrong, but this intuition cannot be captured or explained by appealing to the constitutive harm of risk *viz.* frustration one's preferences.

A second general problem is that the preference account generates counterintuitive verdicts in at least two kinds of cases: one, cases where agents have actual preferences *for* being subject to (higher) risks of harm; and two, cases where an agent's actual preference for being subject to risk is based on a mistaken or false belief that the risked harm or the exposure to risk itself is good for them. As an illustration of the first case, think of a risk-loving passenger who strongly prefers to fly in defective planes because the risk involved gives her an adrenaline rush. As an illustration of the second case, think of a risk-seeking passenger who prefers to fly in a defective plane because she wrongly or mistakenly believes that facing a higher chance of death will bring her good fortune. In these cases, risk itself fails to constitute any harm *viz.* frustration of the risk-bearer's preferences. Accordingly, because the risk is not harmful, the preference account delivers the verdict that Bill's risky conduct is not wrong, contrary to our intuition that Bill in fact acts wrongly.

But perhaps we can put these problem cases aside. The preference account only needs to deliver the correct verdict in *some* cases to affirm that risk is constitutively harmful, and thereby wrong. Let's reconsider the airplane example. Suppose you are one of the passengers.

¹⁴ Some theorists argue that one must exist in order to have preferences. C.f. Bradley 2016.

Suppose I inform you that because of Bill's negligence, you belong to a reference class consisting of all those individuals who fly in defective airplanes and for whom the relative frequency of dying in a plane crash is 50%. Had Bill not acted carelessly, you would have travelled in a safer airplane and faced much lower risk. Is Bill's risk imposition constitutively harmful and thereby wrong?

On the preference account, the answer depends on what you prefer and whether Bill's actions frustrate *that* preference. Suppose you respond by saying that you prefer not to face 50% risk of dying. Let's label this as your preference "*P*". If we imagined a harm-o-meter that keeps record of when risk is constitutively harmful, the meter registers the harm of risk imposition at *t*. Why? Because Bill's action brings about a state of affair at *t*, namely, exposing you to 50% risk of dying, that frustrate *P* at *t*. The preference account thus delivers the verdict that Bill's risk imposition is wrong by virtue of harming you viz. desire frustration.¹⁵

This judgement, however, can be questioned. It seems to hold true only on a particular description of the state of affairs (the risk itself) that obtains because of what Bill does at *t*. There are, however, other ways of describing the state of affairs. As is well known, on a relative frequency interpretation of objective risk, the risk that one imposes upon another is defined relative to whatever reference class one selects (Perry 2007; Oberdiek 2017). It is thus presumed that there is no singular 'correct' risk that one is exposed to at a given time, but rather multiple risks, each defined relative to a certain, specific reference class. Accordingly, since a particular individual can belong to many such reference classes, there are, at least in principle many different characterizations or descriptions of risk that one is exposed to.¹⁶

Suppose I tell you that because of Bill's risky action, you also belong to a reference class consisting of all those individuals who fly in airplanes *on a Thursday* and for whom the relative frequency of dying is 5%. Accordingly, the harm-o-meter does *not* register that risk constitutes harm at *t*. Why? Because Bill's action brings about a state of affair at *t*, namely, exposing you to 5% risk of dying, that does not frustrate *P* at *t*. The preference account thus delivers the verdict that Bill's risk imposition is *not* wrong insofar it is not constitutively harmful viz. desire frustration.

Accordingly, the status of risk being constitutively harmful on your harm-o-meter changes as we go about redescribing risk at *t*. It changes with whether or not the description of the state of affair (risk) that obtains at *t* corresponds with the frustration or satisfaction of your individual preference *P*. Since both (and many other) simultaneous characterizations of risk are plausible (that is, the risk you face at *t* can be *both* 50% or 5% or something else depending on our chosen reference class), the preference account entails that risk can be both constitutively harmful and not harmful viz. frustration of your desire at time *P* at *t*.

The upshot is that this contrasting judgement regarding the harm suffered by you results in at least two contrasting judgements of the wrongness of Bill's risk imposition. Since Bill does not frustrate your preference *P* on one description of risk, what he does is not *pro tanto* wrong. At the same time, what he does is *pro tanto* wrong on a distinct description of the risk, even though Bill happens to perform exactly the *same* action. This is an odd consequence. However, it is an inevitable one for the preference account.

The reason is as follows. On an actualist preferentialist theory that Finkelstein relies on, increments or diminishment of our well-being "consists in achieving an optimal overall match

¹⁵ Note that I use the word 'desire' as a synonym for 'preferences'.

¹⁶ There is no strict criteria for characterizing risk one way or the other for it can vary with context and purpose. This is the well-known reference class problem (Reichenbach 1949, Hájek 2007).

between the way things are and the way we want them to be, whatever we may happen to want” (Lukas, 2010, p.2). In the case of risk, the way things are (that is, the risk you face) can be described in a multitude of ways: the fact that the risk to you at t is 5% relative to particular reference class is a mind-independent feature or state of the world described in one way; but so is the fact that at t , the risk is 10% relative to a different reference class.

Accordingly, whether or not the risk one faces is constitutively harmful depends on, and varies with whenever there is a one-to-one optimal match or correspondence between the object of one’s risk preference (P) and the concurrent obtaining of a particular state of affair (risk one faces) described in a certain way. This entails the possibility that risk can be both harmful and not harmful at the same time. This makes the preference account implausible for at least two reasons.

First, it seems to vindicate an odd conception of constitutive harm of risk such that whether it obtains (as well as the degree or extent to which it obtains) depends on the way *we* choose to describe risk and the corresponding frustration of an individual’s specific preference for that risk. The oddness comes to light when we consider a paradigmatic case of ‘harm’, such as causing someone pain. If I stab my friend’s leg with a stone, whether or not he suffers pain is not in any way relative to my description of what I do.

Second, the account seems to controversially rely on the idea that individuals have very specific, fine-grained preferences corresponding to which reference class they are exposed to, and that there is a natural or obvious way of individuating such preferences.¹⁷ It’s also not clear whether individuals actually have such fine-grained preferences in the first place, and whether these can in fact be individuated and measured in any clear or systematic manner. Thus, it seems that we cannot unproblematically locate the constitutive harm of risk in frustration of one’s preference against specific risk.

Moving forward, there is one final candidate left to be explored. It seems *prima facie* plausible that if the passengers prefer not to be harmed, say, incur an injury at any given moment in time, then being exposed to the risk of getting injured frustrates *this* preference. Let’s label it as preference “ P^* ”. The object of P^* is distinct from the preference P considered earlier – the intentional content of former corresponds to one’s desire not to be exposed to outcome-harm *per se*, whereas the intentional content of latter corresponds to one’s desire not to be exposed to risk *per se*.

When Bill imposes the risk of a plane crash at t , it is both indeterminate and beyond Bill’s control whether or not the object of P^* , namely, outcome-harm will in fact obtain at t or some later time. Thus, we could say that *ex ante*, the imposition of risk itself *entails* or *constitutes* frustration of P^* , even if the risk eventually doesn’t materialize *ex post*.¹⁸ Accordingly, perhaps Bill’s risk imposition at t is constitutively harmful, and thereby wrong because risk frustrates P^* at t . There are, however, two problems with this proposal.

First, on commonly accepted accounts of preferentialist theories, the harm (that is, reduction in well-being) of an episode or an instance of preference-frustration is a product of two things: the intensity of the desire in question (more intense a desire, more disvaluable its frustration) and the duration of the episode (Heathwood 2006; Skow 2017; Sarch 2013). Now compare two cases involving frustration of P^* . In case one, Bill exposes you to a very high risk of injury, say 50% for one hour. In case two, Bill exposes you to a very low risk of injury, say

¹⁷ I thank an anonymous reviewer for pointing this out.

¹⁸ Note that this is different from saying that Bill’s risk imposition at t entails a *risk* of frustration of P^* such that actual frustration of P^* is conditional on this risk materializing.

1% for one hour. Suppose the intensity of P^* , namely, your desire not to be injured remains constant in both cases, say, the intensity is 10 units. Furthermore, suppose only the duration of the episode of desire frustration varies in both.

Here is what an actualist preferentialist theory will say about the constitutive harm of risk: other things being equal, you suffer 10 units of harm (product of intensity (10) and duration (one hour)) in both cases. Accordingly, *if* it is true that Bill's risk imposition is constitutively harmful when and *because* he frustrates P^* , then the preference account delivers the verdict that Bill wrongs you to the same extent when he exposes you to either 1% risk of injury or 50%. This, however, is counterintuitive. Intuitively, *if* risk is constitutively harmful viz. frustration of one's desire not to be harmed, *then* it is worse to expose someone to a *very* high risk of injury for an hour than to a *very* low risk for an hour, other things being equal. This is because higher risks frustrate P^* more severely (that is, more reduction in well-being) and lower risks frustrate P^* less so, all other things being equal. The preference account, however, cannot capture this.

The second problem goes as follows. Again, compare two cases involving frustration of P^* . In one case, Bill exposes you to a very high risk of injury, say 50%, but the risk fails to materialize after an hour. In case two, Bill exposes you to a very high risk of injury, say 50%, but this time, the risk materializes within a minute. Keeping intensity of P^* constant, say 10 in both cases, the actualist preferentialist theory says that on *ex post* judgement of the case, the harm of risk itself in case two is *less* than in case one at a given time-slice: the product of intensity and one hour long frustration of P^* in case one is more than product of 10 and one minute of frustration of P^* in case two. This, however, runs contrary to our intuitive judgement. The degree of wrongness of Bill's risk imposition is more if his actions actually result in outcome-harm as compared to a case where he merely imposes *pure* risk of harm. The preference account, however, cannot capture this.

I conclude that these considerations count against the success of preference account in locating the wrong of imposing risk in risk being constitutively harmful.¹⁹ Perhaps a different account works better.

4 Constitutive Claim: Risk's Impact on Interests

Some theorists have proposed appealing to a distinct currency of risk-harm. In particular, they assume a Feinbergian (1987) conception of harm and locate the constitutive harm of risk in how risk impacts risk-bearer's interests rather than her well-being *per se* (Oberdiek 2012, 2017; Placani 2016).²⁰ This shift in focus avoids many problems the preference account faces. Interests, unlike preferences, are just some things that one has a *stake in* (Feinberg, 1987, p.42). For instance, what promotes an agent's interest in health is good for her, even if she mistakenly believes the contrary, and even if she really desires ill health. As Feinberg argues, a

¹⁹ Finkelstein also takes the harm of risk as necessary but not sufficient to establish a claim of liability or compensation (Finkelstein 2003, p.963). However, if risk itself is not harmful as I've argued *contra* Finkelstein, then it raises further question of what, if not the harm of risk itself, grounds liability in cases of wrongful risk impositions.

²⁰ Feinberg (1987, p.32-7) differentiates between a non-normative, legal notion of harm as a set-back to an individual's core interests, and a normative notion of harm as wrongful set-backs to interests that involves violation of rights (p.36). Oberdiek (2009, p.380) and Placani (2016, p.92) appeal to the former notion.

mere desire, no matter how genuine or intense, does not establish an interest in the sense required for establishing a harmed condition on his framework (p.56).

Thus, by identifying the interest(s) that can be affected and the manner in which they are affected by risk itself, we can anchor the wrong of imposing risks in the following way: other things being equal, an agent is made worse-off, and thereby wronged, when and because risk sets-back one or some of her core, legitimate interests.

So, what are these interests? According to John Oberdiek (2009, 2012, 2017), a plausible candidate is autonomy. He notes, "...while imposing risk does not involve material harm, like the injuries suffered in a car accident, it nevertheless can constitute a setback to a non-material autonomy interest of a certain kind" (2017, p.71).²¹

In thinking about autonomy, Oberdiek has Joseph Raz's (1986) conception in mind. On Raz's account, a person is autonomous if she can be considered the author of her own life. Autonomous living has multiple necessary preconditions, such as having the basic mental capacities to form and execute intentions, freedom from coercion or manipulation, being able to chart one's own course through life, and having plentiful or an adequate range of options to choose from.

According to the autonomy account, in some cases, pure risk imposition constitutes harm, and thereby wrongs the agent when and because it hurts a certain pre-condition of autonomy.²² This pre-condition is having an adequate range of acceptable options to choose from. In explicating how risk affects this pre-condition, Oberdiek (2017) asks us to think of imposing risks as laying traps. Pure risk imposition, the argument goes, has the same effect for risk-bearers as a trap: "the trap takes away the option, or more accurately renders unacceptable the exercise of the option, of stepping where the trap has been set" (p.86). If enough traps are laid, one's autonomy would be severely diminished.

Similarly, in subjecting someone to risk, we foreclose someone's safe options. Lack of safe options has a constraining effect: it narrows one's range of acceptable, choice-worthy options. Accordingly, sometimes, exposing someone to risk *entails* a setback or diminishment of one's autonomy. Since setting-back one's autonomy interest is to make someone worse-off (*following* Feinberg), Bill's risk imposition is constitutively harmful, and thereby wrong. Is this verdict correct?

Despite its initial plausibility, I argue that the claim is found wanting. To begin with, consider Oberdiek's appeal to what he calls 'non-material autonomy interest'. On Feinberg's (1987, p.33) framework, when we speak of interests, we usually have in mind things which autonomous agents are invested in. These interests, or more accurately, the things these interests are *in* are related either to an agent's ulterior personal goals and aims (such as writing a novel, living a flourishing life, maintaining physical health, etc.) or related to conditions and goods that make it possible for agents to achieve their ulterior goals and aims (such as bodily integrity, emotional stability, physical health and vigor, minimal income, etc.). Feinberg labels former as one's *ulterior interests*, and the latter as one's *welfare interests* (p.38–9). On his

²¹ A related account offered by Placani (2016) appeals to risk-bearer's dignity interests. On her view, imposing risks on people with the intention of harming them can sometimes undermine their status as people worthy of respect. The cases discussed here, such as Bill's reckless drunk driving or negligence in the pilot example do not involve this intention. Accordingly, Placani's account, I claim, would not treat them as cases of harmful risking.

²² Parr and Slavny (2019), p.4) also discuss Oberdiek's account but treat it as an alternative to Harm View. However, Oberdiek himself defends his account as a "different kind of harm-based account of the moral significance of risking" (2012, p.350)"

account of harm, only affecting the latter counts as harm and only sometimes, causing such harm is wrongful (p.39).

Autonomy – understood here in Razian terms (*following* Oberdiek) - is not included in Feinberg's list of welfare interest. One might think that regardless, autonomy is a "basic pre-requisite" for achieving one's ulterior interests and thus earns a rightful place on this list. If so, setting-back one's autonomy interest by exposing them to risk constitutes harm and is thereby wrong. However, there are reasons to think otherwise.

On Feinberg's framework, welfare interests are interests in *conditions* that are generalized means for the achievement of more ultimate goals, namely, one's ulterior interests. They require protection and preservation because without them, a person is simply lost. However, Feinberg makes clear that welfare interests are to be treated only as "minimal but non-ultimate goods" that are "basic requisites of a man's well-being, but by no means the whole of that well-being itself." (p.38). Accordingly, if autonomy is a welfare interest, then is a relatively trivial good that is necessary but insufficient for a good life.

In contrast, on Raz's framework, the ideal of autonomy is construed *not* as a means to a good life, but as a kind of achievement that is an essential element of the good life itself. For Raz, an autonomous life is of ultimate, intrinsic value rather than instrumentally valuable for achieving one's ulterior interior. As Raz himself notes, "if having an autonomous life is an ultimate value, then having a sufficient range of acceptable options is of intrinsic value, for it is constitutive of an autonomous life that it is lived in circumstances where acceptable alternatives are present." (p.205).²³

Respecting people's autonomy requires that we give due considerations to all their interests that promote the preconditions for their autonomy or achieving an autonomous life. Autonomy, then, is *not* an interest in and of itself. Rather, it acts as a constraint on setting-back interests of agents that are pre-requisites for leading an autonomous life.²⁴ But if we treat autonomy (in the Razian sense) as a welfare interest (in the Feinbergian sense) as Oberdiek does, then it is unclear how we can reconcile the Razian idea of autonomy having ultimate value with the Feinbergian idea of autonomy as not having ultimate, but typically only instrumental value.²⁵ The two ideas are at odds with each other.

Thus, if autonomy as an interest falls out of the Razian framework, then this has an important implication for Oberdiek's proposal: it follows that risk is *not* constitutively harmful. Why? If the constitutive harm of risk consists in setting-back of an interest, and if autonomy (in the Razian sense) is *not* a welfare interest in and of itself (in the Feinbergian sense) for the reasons discussed here, then risk is *not* constitutively harmful by virtue of setting-back one's autonomy *interest* as the autonomy account claims.

Perhaps we can simply drop the language of interests and autonomy account's commitment to Feinberg's notion of harm, and reformulate Oberdiek's proposal as follows: sometimes, imposing risk constitutes harm, and is thereby wrong, when and because risk negatively impacts one's autonomy. Is this verdict correct?

Despite this reformulation, the autonomy account is open to a number of objections. First, consider a general problem. Reliance on autonomy implies that risk imposition is harmful only for those who, at the time, are leading autonomous lives or have the potential to do so. Suppose

²³ As Raz further notes, autonomy is of value to people living in certain societies *only* and hence, not everyone has an interest in personal autonomy. C.f. footnote 103, p.189 in *The Morality of Freedom*.

²⁴ Most theorists who defend a will-based theory of rights also hold this view.

²⁵ C.f. Feinberg (1987, p.57–61) in *Harms as Setbacks to Interest* where he develops this point.

that Bill is laying risky traps on the road and a baby crawls towards them. If imposing risk is wrong because it is harmful viz. diminishing one's autonomy, then Bill does not wrong the baby by virtue of harming her. Similar verdict holds in a case involving those facing autonomy-deficits, like prisoners, or agents with not-yet developed rational capacities necessary for autonomous living.

Thus, in these cases, the constitutive harm of risk viz. autonomy-diminishment fails to be the correct explanation of what makes imposing risk wrong. This failure is common to another type of where wrongful risk imposition forecloses or removes one's option but fails to be constitutively harmful. Recall Oberdieks' position that sometimes risk diminishes autonomy by affecting one's *adequate* range of acceptable or safe options that a given person could exercise (2012, p.385). But what constitutes an 'adequate' range? Raz (1986) notes that "to satisfy the conditions of the adequacy of the range of options, the options available must differ in respects which may rationally affect choice...or else the requirements of variety which is a precondition of the adequacy of options will not be met." (p.398).

Accordingly, having an adequate range of option requires having a set of diverse, choice-worthy options to choose from. This constitutes a benefit insofar it fosters the condition for "determining the course of one's life oneself by realizing some options and blocking others" (Raibley 2018 p.77). Conversely, lacking a range of diverse, valuable options to choose from, like a man stuck in a pit with the only option of eating and sleeping is harmful insofar it is detrimental to living an autonomous life. As Raz notes, "not number but variety matters" (1986, p.375).

Thus, if risk itself is constitutively harmful and thereby wrong, then it is so only when it affects the *variety* of one's safe, choice worthy options. However, there are cases wherein wrongly imposing risk on others forecloses range of similar options, but fails to have any negative affect on the *variety* of one's option set. Suppose that Bill lays a risky trap on my walking path. Next morning when I get to the path, I notice the trap and decide to take one of the other different paths that are still open to me. While Bill's risk imposition may have constrained my choice by removing one option, the important thing is that I still have *some* variety of safe options left to choose from. As long as risk does not intervene with *this*, my autonomy is not interfered with in a way that makes it constitutively harmful. Accordingly, even if Bill's imposition of risk in this case is wrong, its wrongness is not explained by the fact that risk is constitutively harmful viz. autonomy-diminishment.

One might object that the autonomy account only needs to show that in *some* cases, imposing risk is constitutively harmful, and when it is, it is wrong. There is, however, still one reason to doubt whether it can do this successfully. Return to the argumentative core of Oberdiek's (2012) account, namely, the analogy that imposing risk is *like* laying traps.²⁶ The analogy drives home the point that just as laying a trap takes away safe options by creating the possibility that someone will get caught in it, imposing risk also takes away safe options by creating the mere *possibility* that someone will suffer some material harm (p.352). The probability is not relevant to whether risk affects one's autonomy. As Oberdiek (2012) himself notes, "while risk is usually (and rightly) regarded as a probability of harm, what is relevant about risk here is an entailment of this, namely, that a risk imposition also connotes the possibility of harm" (p.352).

²⁶ Note that Oberdiek argues that imposing risk is *like* laying traps, rather than arguing that laying trap is an *instance* of risk imposition.

Accordingly, the explanation for *what* makes risk constitutively harmful viz. diminishing one's autonomy is that sometimes risk constitutes the mere *possibility* of suffering outcome-harm, and when it does, the mere possibility suffices for risk to be harmful. Unfortunately, this way of construing the constitutive harm of risk has an implausible consequence.

Compare three cases. In case one, Bill sets up a trap in front of your door that works 1 out of 2 times, thereby exposing you to 50% risk that you will get caught in it *if* you step into the trap. In case two, everything is the same expect that now the trap works 1 out of 20 times, thereby exposing you to 5% risk. In case three, again everything is the same expect that now the trap works 1 out of 200 times, thereby exposing you to 0.5% risk. In each case, the risk you face constitutes or connotes the *possibility* of harm even though the *probability* differs. As there is no difference between diminishing one's autonomy when we expose others to very high or very low risk of harm - as only the *possibility* of future harm matters for reasons noted above - the autonomy account delivers the verdict that your autonomy is diminished to the same degree if you are exposed to 50% risk or 0.5% or anything in between.

Accordingly, risk is harmful in the same way and to the same degree in every case. This entails that the degree of wrongness involved in imposing risk in *all* the three cases of risk is also at par.²⁷ This, however, is counterintuitive. Intuitively, it seems that *if* one is harmed by risk imposition viz. autonomy-dimishment, then one is harmed more if the risk to self is 50% than when it is, say 5%. However, the autonomy account cannot capture this intuition. Moreover, by rendering *any* degree of risk as constitutively harmful, the autonomy account seems to treat our everyday, mundane actions as instances of harm insofar all our actions carry some risk for others. In doing so, the account risks trivialises the moral significance of harming much in the same way we encountered earlier with the preference account.

That the harm of risk is rendered trivial becomes evident when comparing a case of pure risk with a case where the risk materializes. Suppose again that Bill sets up a trap in front of your door. In one case, you face a 5% risk of getting trapped but luckily nothing happens. In another case, the risk materializes and you get trapped. In the former case, the autonomy account delivers the verdict that you are harmed because the *risk* of getting trapped diminishes your autonomy: your safe options of stepping out it foreclosed. The account delivers the *same* verdict for the latter case where you are *actually* trapped: your safe options to move around are foreclosed. Despite the difference in which of your option(s) are affected in either case, the outcome with respect to suffering the harm of autonomy-dimishment is the same in both cases. This, however, is problematic because it blurs the intuitive normative difference between wrongful autonomy -diminishment when suffers the outcome-harm, and when someone is exposed to mere *risk* of that harm.

I conclude that these considerations count against the success of autonomy account in locating the wrong of imposing pure risk in risk being constitutively harmful. If defenders of the Harm View still want to explain what makes pure risk imposition wrong in terms of risk itself being harmful, they must locate the harm of risk elsewhere. This brings us to the Contingent Claim.

²⁷ *Mutatis Mutandis* for any other instance of autonomy-diminishing risk imposition.

5 The Contingent Claim

The Contingent Claim holds that pure risk imposition is *pro tanto* wrong when and because risk has harmful consequences for the risk-bearer. Accordingly, that risk itself is harmful and thereby wrong depends on what risk *causes* rather than what it itself *constitutes*. The Contingent Claim is thus immune to the various problems that the Constitutive Claim faces.

Theorists often appeal to psychological distress such as fear and anxiety in spelling out the Contingent Claim (Wolff and De-Shalit 2007, Arneson 2010, Cripps 2013). Suppose that you see Bill driving recklessly towards you. Luckily, you notice in time and jump out of the way. Your quick reaction prevents the risk of an accident from materializing for you. Nevertheless, as you experienced unpleasant levels of stress and anxiety, Bill's risk imposition is *pro tanto* wrong because risk has these harmful effects for you. In some cases, it can even lead to further harmful effects. For instance, the fear induced by being exposed to risk may force you to go out of your way to avoid the risk by taking costly or difficult precautionary steps (Wolff and De-Shalit 2007).

It is very plausible that experiencing these effects is detrimental to your well-being. The Harm View, then, successfully generates a harm-based explanation for why imposing risks is wrong in cases where risk itself has harmful effects. However, this also means that the Harm View is limited in its explanatory scope. As the name itself suggests, the Contingent Claim only captures cases where the relevant, contingent effect obtains. For instance, suppose that unbeknownst to you, Bill drives recklessly towards you while you are walking blind-folded on the street. Consequently, this time you do not experience any psychological distress. Intuitively, Bill wrongs you by subjecting you to risk, independently of the relevant harmful effects obtaining. The Harm View, however, cannot offer us an explanation for what makes Bill's conduct wrong.

Moreover, in cases where the relevant, harmful effects do obtain - whatever these effects may be - it is not always clear whether risk *itself* is one of the necessary (but insufficient) causes of the harmful effect. For instance, suppose that as a matter of fact, the chances that Bill injures you by driving around recklessly is one in a thousand. When you see him driving towards you, you experience a lot of distress and fear because you overestimate the risk of him crashing into you. You think that Bill is in fact exposing you to a one in two chance of a serious accident.

Despite Bill's risk imposition resulting in harmful effects for you, it is at least plausible, if not necessary, that these effects are caused by your heightened or exaggerated perception and/or evaluation of *riskiness* of Bill's driving. Or perhaps these harmful effects could be caused by your expectation or anticipation of getting into an accident with him. Or it might be that your distress is facilitated or mediated by your awareness or knowledge that you are being exposed to some non-negligible risk that may materialize for you. And so on.

In such scenarios, if risk itself is not one of the necessary (but insufficient) causes of the harmful effect, then Bill's risky conduct is not wrong by virtue of the risk itself being contingently harmful, even though it is bad for you to experience these effects. Bill wrongs you by subjecting you to even a one in a thousand risk of crashing into you, however, the wrongness is not explained by risk *itself* being contingently harmful.

This does not necessarily count as a strike against the Harm View, insofar it correctly captures some cases where it is risk itself that has harmful effects for risk-bearers. It only means that the Harm View is insufficient as a general, exhaustive account of what makes imposing risks wrong for two reasons. First, in cases of wrongful risk imposition where

contingent harmful effects are present, if risk *itself* is not one of the necessary (but not sufficient) causes of these effects, then risk *itself* is not harmful and thereby wrong. Second, that risk itself is contingently harmful and thereby wrong is only true for a narrow range of cases. This means that if we want to capture the wrongness of a wider range of risk imposition cases, we need to look beyond the Harm View.

6 Conclusions

In this paper, I have engaged in a detailed discussion of a prominent view in the ethics of risk imposition, namely the Harm View. I've argued that the Harm View is plausible only on the Contingent Claim, but fails on the Constitutive Claim. In discussing the Constitutive Claim, I've argued that the preference and autonomy accounts as construed by Finkelstein (2003) and Oberdiek (2017), respectively fail to show that risk itself is constitutively harmful, and thereby wrong. In vindicating the idea that risk itself is constitutively harmful, both accounts are found guilty of either trivializing or undermining the moral significance of risk, or admit to having counter-intuitive implications in cases where risks materialize. In discussing the Contingent Claim, I've argued that risk itself is contingently harmful and thereby wrong only in a narrow range of cases. This makes the Harm View explanatorily limited in scope, thereby undermining its plausibility as a general, exhaustive account of what makes pure imposition wrong.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arneson R (2010) Book review: disadvantage, capability, commensurability, and policy. *Polit Philos Econ* 9(3): 339–357
- Bradley B (2016) Well-Being at a Time. *Philosophic Exchange* 45(1).
- Cripps E (2013) *Climate change and the moral agent*, 1st edn. Oxford University Press, Oxford, pp 27–57
- Feinberg J (1987) *The moral limits of the criminal law volume 1: Harm to Others*, 1st edition. Oxford University Press, New York
- Ferretti M (2016) Risk imposition and freedom. *Polit Philos Econ* 15(3):261–279
- Finkelstein C (2003) Is Risk a Harm? *Univ Pa Law Rev* 151(3):963–1001
- Gardner M (2017) On the strength of the reason against harming. *J Moral Philos* 14(1):73–87
- Gillies D (2000) *Philosophical theories of probability*, 1st edn. Routledge, London
- Hájek A (2007) The reference class problem is your problem too. *Synthese* 156(3):563–585
- Hanser M (2013) "Harm," in H. LaFollette (ed.), *International Encyclopedia of Ethics*. Oxford: Blackwell Publishing, pp. 2299–2307
- Hayenhjelm M, Wolff J (2011) The moral problem of risk impositions: a survey of the literature. *Eur J Philos* 20: E26–E51
- Heathwood C (2006) Desire Satisfactionism and hedonism. *Philos Stud* 128:539–563

- Herington J (2017) Climate-related insecurity, loss and damage. *Ethics Policy Environ* 20(2):184–194
- Herington J (2019) The contribution of security to well-being. *J Ethics Soc Philos* 14(3):179–203
- Hurd H (1996) The deontology of negligence. *Boston Univ Law Rev* 76:247–272
- Lazar S (2019) Risky Killing. *J Moral Philos* 16(1):1–26
- Lukas M (2010) Desire satisfactionism and the problem of irrelevant desires. *J Ethics Soc Phil* 4(2)
- Mill JS (1859) *On liberty*. Parker, London
- Oberdiek J (2009) Towards a right against risking. *Law Philos* 28(4):367–392
- Oberdiek J (2012) The moral significance of risking. *Legal Theory* 18(3):339–356
- Oberdiek J (2017) *Imposing risk: a normative framework*, 1st edn. Oxford Legal Philosophy
- Parfit D (1984) *Reasons and persons*. Clarendon Press, Oxford
- Parr T, Slavny A (2019) What's wrong with risk? *Thought J Philos* 8(2):76–85
- Perry S (1997) In: D. Owen, ed., *Philosophical Foundations of Tort Law*, 1st ed. Oxford: Clarendon Press, pp 322–346
- Perry S (2007) Risk, harms, interests, and rights. In: Lewens T (ed) *Risk: philosophical perspectives*, 1st edn. Routledge, New York, pp 190–201
- Perry S (2014) Torts, rights and risk. In: Oberdiek J (ed) *Philosophical foundations of the law of torts*, 1st edn. Oxford University Press, Oxford, pp 38–64
- Placani A (2016) When the risk of harm harms. *Law Philos* 36(1):77–100
- Raibley J (2018) Paternalism and Well-Being. In: Grill K, Hanna J (eds) *The Routledge handbook of the philosophy of paternalism*, 1st edn. Routledge, London, pp 69–84
- Raz J (1986) *The morality of freedom*. Oxford University Press, Oxford
- Reichenbach H (1949) *The theory of probability*. University of California Press
- Roser D (2017) The irrelevance of the risk-uncertainty distinction. *Sci Eng Ethics* 23(5):1387–1407
- Sarch A (2013) Desire Satisfactionism and time. *Utilitas* 25(2):221–245
- Shiffrin SV (2012) Harm and its moral significance. *Legal Theory* 18(03):357–398
- Skow B (2017) Preferentism and the paradox of desire. *J Ethics Soc Philos* 3(3):1–17
- Tadros V (2016) *Wrongs and crimes*, 1st edn. Oxford University Press, Oxford, pp 175–200
- Thomson J (1986) *Rights, restitution, and risk: essays in moral theory*, 1st edn. Harvard University Press, USA, pp 173–191
- Wolff J, De-Shalit A (2007) *Disadvantage*. Oxford University Press, Oxford

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.