UCLA
UCLA Public Law & Legal Theory Series

Title
HARM AND ITS MORAL SIGNIFICANCE

Permalink
https://escholarship.org/uc/item/2dz5b9rk

Journal
Legal Theory, 18(3)

ISSN
1352-3252 1469-8048

Author
Shiffrin, Seana

Publication Date
2012-08-16

DOI
10.1017/S1352325212000080

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

*Legal Theory*, **18** (2012), 357–398.
© Cambridge University Press 2012 0361-6843/12 $15.00 + 00
doi:10.1017/S1352325212000080

# HARM AND ITS MORAL SIGNIFICANCE

Seana Valentine Shiffrin*
*Department of Philosophy and School of Law, University of California, Los Angeles*

Standard, familiar models portray harms and benefits as symmetrical. Usually, harm is portrayed as involving a worsening of one's situation, and benefits as involving an improvement. Yet morally, the aversion, prevention, and relief of harms seem, at least presumptively, to matter more than the provision, protection, and maintenance of comparable and often greater benefits. Standard models of harms and benefits have difficulty acknowledging this priority, much less explaining it. They also fail to identify harm accurately and reliably. In this paper, I develop these problems, argue that we should reconsider our commitment to the standard models, and then merely gesture at the direction in which we might locate a superior approach, one that better accounts for the moral significance of harm and its relation to autonomy rights.

## I. INTRODUCTION

The concept of harm plays a central role in moral, political, and legal theory. Restrictions on harming ground key deontological prohibitions; harm helps to set the limits of the liberty principle; harm often serves as a key element of legal liability; and the goals of avoiding, preventing, and alleviating harm occupy a major part of the moral agenda for those who aim to aid others. Yet harm's centrality is underexamined and still poorly understood. Worse still, standard accounts of what harms and benefits are conflict with our moral intuitions about the significance of harm.

*Largely, I wrote this paper more than a decade ago as an extension of some arguments made in Seana Valentine Shiffrin, *Wrongful Life, Procreative Responsibility, and the Significance of Harm*, 5 LEGAL THEORY 117–148 (1999). I put it aside in dissatisfaction. I remain dissatisfied. Nonetheless, despite my lingering doubts, I have not found sufficient cause to reject its main ideas either. The manuscript has circulated a little over the years. Because some have cited (and criticized) it, Gregory Keating convinced me it might be useful if it were publicly available, despite its flaws. I publish it with hesitation, however. I wish it were better, clearer, and shorter. I am grateful to a number of people and audiences for patient, insightful help and critical encouragement. I owe particular thanks to Janet Broughton, Tyler Burge, Meir Dan-Cohen, Hannah Ginsborg, Matthew Hanser, Barbara Herman, Frances Kamm, Sanford Kadish, Gregory Keating, Herbert Morris, Michael Otsuka, Samuel Scheffler, Steven Shiffrin, Terry Stedman, Matthew Strawbridge, Judith Jarvis Thomson, J. David Velleman, Daniel Warren, and Stephen White.



https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

Standard, familiar models portray harms and benefits as symmetrical. Usually, harm is portrayed as involving a worsening of one's situation, and benefits as involving an improvement. Yet morally, the aversion, prevention, and relief of harms seem, at least presumptively, to matter more than the provision, protection, and maintenance of comparable and often greater benefits.[1] Standard models of harms and benefits have difficulty acknowledging this priority, much less explaining it. They also fail to identify harm accurately and reliably. In this paper, I develop these problems, argue that we should reconsider our commitment to the standard models, and then merely gesture at the direction in which we might locate a superior approach, one that better accounts for the moral significance of harm and its relation to autonomy.[2]

The argument, such as it is, proceeds in four stages. First, I identify some preliminary desiderata for a satisfactory theory of harm and articulate some features of harm's priority. Second, I argue that standard accounts of harm cannot achieve these desiderata. Third, I discuss some difficulties squaring certain prominent understandings of autonomy rights with the relation of autonomy rights to harm, whether conceived of in a standard way or otherwise. I suggest that these difficulties give us reason to think of autonomy in a somewhat different way and that they generate further desiderata for an apt theory of harm. Fourth, these considerations in turn point toward the outlines for an account of harm that could make sense of harm's moral significance more successfully and that could support a more integrated account of the relationship between autonomy rights and harm's significance. My suggestions about a better positive account of autonomy rights, harm, and their relationship remain tentative, speculative, and general. I offer them more in the spirit of thoughts to open (or further) a discussion than as elements of a concrete, precise proposal.

## II. DESIDERATA

I begin by articulating two preliminary desiderata for a theory of harm and by making plain some aspects of the paper's methodological orientation.

First, a theory of harm to individuals should identify harm in a way that captures most of our central, intuitive judgments about what counts as harm, while supplying some unifying understanding of why all the conditions we

1. This asymmetry figures prominently in most nonconsequentialist moral theories but also figures in sophisticated consequentialist accounts that give special weight to harm minimization.

2. Although these issues are closely related to the significant role harm plays in law, I focus here only upon the moral concept of harm and I do not attempt an account of the proper role it should play in law. Still, I believe that an adequate account of the moral concept should help make sense of the significant role the concept of harm plays in the law. I hasten to add, though, that I do not take a showing of harm to be a necessary predicate for legal liability and hence I do not regard it as one of the obligations of an adequate account of harm that it can support a comprehensive account of legal liability.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

recognize as harms fit together. Second, it should also fit with and support our judgments of how and why harm especially matters, what I refer to as "harm's priority."

Of course, it is unlikely that any single account of harm can unify all the ordinary language uses of the term "harm." "Harm," I freely concede, has a diverse range of ordinary language applications. Probably no account could comfortably and meaningfully accommodate all these uses. "Harm" can be used to signify damage, pure and simple. One may say, for example, that the table is harmed when it is dented by a clumsy mover. And we may say that a plant is harmed when its leaves are stripped or an insect is harmed when its antennae are torn; here we point to an extended sort of damage or interruption of life function. I do not mean to address these sorts of harm. Rather, I mean to try to isolate a core notion of harm to individuals that justifiably occupies a prominent place in normative theory.

To claim that a theory of harm should provide the resources to explain harm's centrality is not necessarily to claim that the concept of harm has a built-in negative moral charge. Killing, for example, is not an essentially normative concept; unlike murder, it does not conceptually convey immorality. Yet killing has a special status in moral theory. It commands a certain level of concern, as opposed to, say, scratching. It routinely requires special justification. If our account of what killing is included a high proportion of relatively minor acts and few that were especially significant, we might wonder why killing in particular calls for special moral notice. Given our strong conviction that killing is presumptively a serious moral matter, we might reasonably decide that our account of killing needed revision without necessarily concluding that killing was an essentially normative concept. Similarly, our sense that harm especially matters is worth taking seriously. Hence it is worth investigating whether this negative valuation of harm could make sense in light of how we characterize harm.

I argue that standard accounts of harm that reduce harm to notions of mere loss, damage, or opportunity costs cannot be squared with harm's priority. For some, I anticipate, the suggestions that either some losses may not be harms or some harms are not losses will be inconceivable. If standard accounts are not reconcilable with the priority on harm, so much the worse for harm's priority.

I take the other tack. Our sense of harm's special importance is worth taking seriously, even if taking it seriously entails that we must reconsider some familiar accounts of harm. For those who opt for the other horn of the dilemma, though, what follows may be better construed as an effort to characterize an aspect of the human condition that properly plays the central moral and prudential role often assigned to harm.

A final methodological note: in characterizing harm and its priority, I focus on the conditions to which individual recipients of harms and benefits are subject and how these states influence our moral considerations about what choices to make and what actions to engage in or omit. Some work on

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

harm shows a tendency, sometimes inexplicit, to reason from facts about the acts or events of harming and benefiting to conclusions about harms and benefits themselves.[3] This can yield misleading results and often eclipses moral phenomena associated with harm. Focusing on the states associated with receiving harms and benefits, whether naturally caused or occasioned by human agency or neglect,[4] may help to disentangle what significance the condition of harm itself (or of suffering, experiencing, or enduring harm, if you prefer) contributes and what is attributable to human agency, causation, and responsibility.[5] Below I suggest that this approach reveals that harm's priority is in some significant respects separate from the distinction between doing and allowing.

Let me begin by elaborating upon the two desiderata of, first, identifying harm in a unified way, and second, doing so in a way that supports its normative significance.

## A. What Counts as Harm

At this early stage, I offer only a first, provisional pass at what an account or theory of harm should include. Briefly, a theory of harm, at least with

3. For some authors, though, this orientation is explicit. *See, e.g.*, Scott Hershovitz, *Two Models of Tort (and Takings)*, 92 VA. L. REV. 1147 (2006), at 1166–1167 (arguing that harms are necessarily the products of wrongs). In a recent set of very thoughtful and inventive articles, Matt Hanser rejects both an act-focused and a state-focused account of harm, advocating an event-focused account of harm. Matthew Hanser, *The Metaphysics of Harm*, 77 PHIL. & PHENOMENOLOGICAL RES. 421 (2008); Matthew Hanser, *Harming and Procreating*, *in* HARMING FUTURE PERSONS 179 (Melinda A. Roberts & David T. Wasserman eds., 2009); Matthew Hanser, *Still More on the Metaphysics of Harm*, 82 PHIL. & PHENOMENOLOGICAL RES. 459 (2011). Judith Jarvis Thomson offers some powerful criticisms of his view as well as some convincing replies to his attack on state-based views; Judith Jarvis Thomson, *More on the Metaphysics of Harm*, 82 PHIL. & PHENOMENOLOGICAL RES. 436 (2011). *See also* Elizabeth Harman, *Harming as Causing Harm*, *in* HARMING FUTURE PERSONS 137 (Melinda A. Roberts & David T. Wasserman eds., 2009), at 148–150.

4. It seems correct to say that a storm, an avalanche, or treacherous waters may bring one harm and that this notion of harm is the same as the notion characterizing the conditions that result from an agent's infliction of similar bodily injury on another. Moreover, it need not have been possible to save the victim, i.e., it need not be an instance of allowing, for naturally inflicted injuries to be harm. Similar reasons to act or to refrain from acting in response to such harms arise whether the harms emanate from natural forces alone or from human action or inaction. I aim to explore the range of overlap and similarity without claiming that the overlap is complete. But even if harm is believed necessarily to involve human agency, whether positive or negative, the paper's main points still hold between doing and allowing harm, on the one hand, and providing and failing to provide or protect benefits, on the other hand. The reason for this must, I think, have something to do with the difference of condition that harms and benefits occasion.

5. I do not mean to rule out the possibility that some harms may be harms only because of their causal origins, whether, e.g., because they were caused by persons under particular circumstances or by persons with specific motives. Further, the severity of some harms may be far worse when they were caused by persons or when they are not accidentally caused but rather caused by, e.g., a negligent or malevolent motive. *See* text accompanying note 43 *infra*. *See also* Stephen Perry, *Harm, History, and Counterfactuals*, 40 SAN DIEGO L. REV. 1283 (2003), at 1294–1295. *But see* John C.P. Goldberg, *Rethinking Injury and Proximate Cause*, 40 SAN DIEGO L. REV. 1315 (2003), at 1322–1323.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

respect to individual human beings, should encompass physical injuries, many physical disabilities, many mental disabilities, some material inabilities, incidents of pain, the failure or ruin of certain sorts of important projects and relationships, some losses, and death. A critical challenge is for the account to capture incidents of pain and the serious frustration of one's central projects as harm while stopping short of counting all desire nonfulfillment as such as harm. Naturally, this goal is connected to the aim of explaining harm's special significance. For if desire nonfulfillment generally counts as harm, regardless of its content and regardless of who possesses it (whether an overindulged utility monster or a neglected waif), then it would seem prima facie implausible that harm in particular could occupy its special place in moral theory.

## B. Harm's Priority

Although many invoke the claim that harm has a special significance, this priority is often underdescribed. It is worth delineating what the priority involves to expose its complexity and to clarify what a theory of harm should aim to accommodate. I describe here two major asymmetries in which harm's priority manifests, although there may be many others. Preliminarily, I rely simply on our intuitive notion of harms and benefits in identifying these asymmetries; below I connect them to more specified notions of harms and benefits.

The first, principal asymmetry arises between the differential reason-generating force and the reason-requiring force exerted by harms and benefits. Generally, other things being equal, harms, harming events, and opportunities to harm are more important morally than benefits, benefiting events, and opportunities to benefit. Harms (and their prospect) have a greater capacity to generate reasons for action and tend to generate stronger reasons for action than benefits (and their prospect). Laying aside moral factors connected to desert, promises, special relationships, and role-generated duties, if a person is (or is about to be) subject to harm, others have *pro tanto* reasons to avoid inflicting it, to prevent it from occurring, or to alleviate it.[6] (Of course, these reasons vary in strength, depending upon

6. Others associate harm with a different moral significance. For example, whereas I focus on the close connection between harms and reasons to alleviate, prevent, and refrain from causing them, Judith Jarvis Thomson draws a tight connection between suffering harm and having a *claim* against someone that they not bring it about. JUDITH JARVIS THOMSON, THE REALM OF RIGHTS (1990), at 228–48. On this point, we differ. Although I think harms provide *reasons* to prevent them, to alleviate them, and to refrain from causing them, I do not believe that those reasons necessarily underwrite *claims* against others. The relevant reasons may not be that strong, for instance. Further, Thomson's connection between harm and *claims against others* may divert our attention away from self-inflicted harm. Differences about the nature of harm's moral significance may in turn generate different judgments about what counts as harm. I mention a few such divergences in what follows, especially note 24 *infra* and Part IV.D.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

the degree of harm and perhaps whether they concern relieving harm or refraining from inflicting it.)

Further, put roughly, the reasons to act generated by harm are stronger than the moral reasons to act generated by comparably sized[7] benefits. Put more precisely, given a benefit and a harm that are comparably sized and both undeserved,[8] stronger moral reasons are generated to refrain from inflicting this harm on a person, to prevent it from occurring (whether by natural or human means), and to alleviate it than are generated to confer this benefit (to that person or another), to facilitate or protect its natural or human conferral, or to prevent its removal if it is already in place.[9]

Moreover, the imposition of harms and benefits calls for the provision of different sorts of justificatory reasons. Generally, the justified imposition of (anything but de minimis) harm on another person requires as a necessary condition that at least one (and perhaps more than one) of the following be true: that the harm be deserved; that its imposition is necessary to avoid greater harm to the recipient or to others; that its imposition be necessary to vindicate another's right; or that the recipient consents. Importantly, consent, where necessary, may not suffice as a justification. Even if the recipient of harm assents, one may not break his nose on a whim, for no reason, or because a lazy thug down the street offered to pay you a hundred dollars to do so. To make the action permissible, the victim's consent and the assailant's action should both be motivated by the reasonable idea that suffering or inflicting the harm will serve a legitimate and suitably important end. On the other hand, to bestow a benefit, the simple assent of the recipient may be needed, especially if the bestowal may alter the feel or character of her life, but generally, this assent will suffice. Other things being equal, one need not have strong justified reasons to believe that the benefit is deserved, necessary to secure greater benefits for the beneficiary or for others, or required to prevent or alleviate harm.[10]

7. The idea of comparably sized benefits and harms is, I concede, more than a little obscure. For my purposes, strict comparability is not essential. We need only have in mind, e.g., the rough comparisons of minor benefits to minor harms.

8. That is, the recipient has not merited and has no special right to the benefit, the harm, or its prevention. The priority on harm is easiest to discern here but arises even with deserved harm, as I discuss in note 33 *infra*.

9. Important moral issues turn on whether and when larger benefits generate stronger reasons for action than lesser harms. Nonetheless, the basic elements of harm's priority do not depend upon a specific stance on this matter. Other asymmetries I discuss, though, suggest that larger benefits do not per se generate stronger reasons than harms. The view that harm generally provides stronger moral reasons to act does not entail that it is always impermissible to pursue a benefit if there is (as there usually is) an opportunity to refrain from, avert, or alleviate harm. Legitimate personal relations, goods of character associated with benevolence, and personal prerogatives may support permissions to pursue some benefits. So this fairly weak characterization of harm's priority can evade the criticism of excessive stringency that are faced by other views closer to negative consequentialism that posit a strong, overriding obligation to avoid or relieve harm. *See, e.g.*, James Griffin, *Is Unhappiness Morally More Important Than Happiness?*, 29 PHIL. Q. 47 (1979).

10. There are complications here with respect to non-anonymous giving, especially when it involves transformative benefits. Such bestowals may substantially alter personal relations,

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

One might consider failures to be benefited to represent the more apt comparison to harm's imposition. Here, too, justificatory asymmetries surface. Failures to benefit may not labor under the same justificatory burden (as with harm) to demonstrate the desert or consent of the potential beneficiary to being passed over; nor does the justification for the omissions depend on its being true that failing to benefit is necessary to prevent further or worse failures to benefit. (If one person is benefited rather than another, though, then fairness concerns may necessitate comparisons of worth, merit, relative position, or type of relation to the benefactor.) Moreover, absent special relational, role, or project-oriented reasons, the general justificatory burden for failures to avert harm is higher than for failures to bestow benefits.

The second asymmetry involves a contrast between what it is permissible to do to avoid or alleviate harms and what it is permissible to do to bestow or protect benefits. There is a fairly broad consensus about what may be called the Rescue Case. Even if a patient is unavailable for consent, so long as there is no strong evidence that his will is opposed, it is permissible to inflict some sorts of necessary lesser harm on him to avert his suffering a significant, substantially greater harm. One may rescue a drowning person from a lake even if the rescue necessitates breaking his arm.[11] A doctor may perform surgery on an unconscious patient to save her life or to prevent or repair severe injury even though the surgery may provoke a painful recovery, damage some limbs or organs, partially disable the patient, incapacitate her for a period, or impose significant risks.[12]

Many read these examples as showing that absent evidence of a patient's contrary will, one may inflict a necessary lesser harm to provide an *overall benefit* to her.[13] This characterization of the Rescue Case is misleading, though. It feeds upon an equivocation about "benefit." In a secondary sense, preventing or alleviating harm is often said to involve benefiting people, and thus the recipient enjoys a "benefit." But the familiar sense in which we say

risking hierarchy and dependence and/or introducing duties of gratitude. These considerations may generate greater justificatory demands for non-anonymously bestowed benefits. Barbara Herman has recently written illuminating work on the subject. Barbara Herman, *Being Helped and Being Grateful: Puzzles about Imperfect Duties, the Ethics of Possession, and the Unity of Morality* (2012) (unpublished manuscript). (Of course, some of these dynamics may also arise with respect to preventions and alleviations of harm.)

11. Joel Feinberg, *Wrongful Life and the Counterfactual Element in Harming*, 4 Soc. Phil. & Pol'y 145 (1986), *reprinted in* Freedom and Fulfillment 3, 27 (1992).

12. There may, of course, be limits to these justifications, depending upon how serious the lesser harm is, how minor the greater harm is, and how close the distance is between the lesser and greater harm. Nonetheless, in a large range of cases, it is permissible to inflict a lesser harm to avert a greater harm.

13. *See, e.g.*, Feinberg, s*upra* note 11; Frances Kamm, Creation and Abortion 173–174 (1992); Jeff McMahan, *Wrongful Life: Paradoxes in the Morality of Causing People to Exist*, *in* Rational Commitment and Social Justice: Essays for Gregory Kavka 208 (J. Coleman & C. Morris eds., 1998); Jeff McMahan, *Cognitive Disability, Misfortune, and Justice*, 25 Phil. & Pub. Aff. 3 (1996); Bernard Williams, *Resenting One's Existence*, *in* Making Sense of Humanity 224 (1995).

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

a benefit is conferred, where this refers to an act of aid, should not be confused with the idea that the recipient's resulting condition is a benefit to her. More important, it should not be confused with the idea that the benefit-bestowing aspect of the saving performs the primary justificatory service for inflicting a harm. The main justification, I believe, is not that assistance is granted but that a greater *harm* is to be averted. A more narrowly tailored understanding of the Rescue Case is that it illustrates that when a patient is unavailable for consultation, it may be permissible to inflict a necessary, lesser harm on her to avert a significant, substantially greater harm.

By contrast, consider a *pure* benefit, by which I mean a benefit that is not also a prevention of, removal from, or alleviation of harm. It is less clear that one may permissibly inflict the same lesser harm to bestow a significantly larger pure benefit.[14] It seems wrong to engage in nonconsensual arm-breaking surgery on a healthy patient that will improve her well-functioning memory or hearing so that it is supersensitive, make her an Olympic-caliber runner, enhance her unmarred appearance, provide her with an ability to imbibe alcohol and fats immoderately and without side effects, or insert into her memory accurate encyclopedic knowledge of Kant's texts.[15] A surgeon who performed this nonconsensual surgery would act impermissibly even if there would be no other opportunity to perform the surgery with consent.[16] And she should compensate for the harm inflicted, unlike in the Rescue Case.

One may object that in the Rescue Case, it is not the greater harm per se that explains the permission. Rather, it is that it is patently clear that the patient would want to avoid the harmful condition, whereas it may not be clear that the patient would want the benefits described. But appeals to the likelihood of appreciation or hypothetical consent do not fully explain the asymmetry. First, there are circularity problems here. One might explain the greater likelihood of hypothetical or *ex post* consent in the Rescue Case as attributable to harm's greater importance, not explanatory of it. Second, the asymmetry still appears even in cases that reflect a similar likelihood of *ex post* appreciation. Even regarding death, there is always some uncertainty that a patient whose will is unknown would really want to avoid it and

14. I discuss this point in Seana Valentine Shiffrin, *Wrongful Life, Procreative Responsibility, and the Significance of Harm*, 5 LEGAL THEORY 117, 121–126 (1999).

15. Frances Kamm has prompted me to note that it would likewise be wrong, without evidence of the patient's will, to perform the harm-inflicting surgery to preserve these benefits. (The fact that the patient had already undergone similar surgery to achieve these conditions, though, might constitute such evidence.) Still, there may be greater reasons to act to preserve an already vested pure benefit, perhaps because of a patient's strong personal identification with it, than to bestow a benefit. But unless the identification is so strong as to make the loss a harm, both nonconsensual surgeries would be impermissible.

16. A weaker version of the asymmetry claims that harming to bestow a benefit requires different, stronger justifications. It may demand that a benefit be of a greater magnitude than the magnitude of the greater harm that justifies lesser harm. Or, it may permit bestowal of the benefit only if the benefactor has a good motive and assumes responsibility and liability for harm caused in the delivery of the benefit.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

the cost of enduring a significant harm. Complementary cases involving benefits of vast sums of money or suitably large conferrals of knowledge may approximate this degree of uncertainty. Even so, it is still easier to justify imposing the lesser harm to avoid a patient's death than to bestow the benefit.[17]

Notice that an important, subordinate asymmetry surfaces here between the assessment of third-party impositions of harm for benefit and impositions that are initiated by the patient herself. While it may be wrong to impose a significant harm on someone to deliver her a greater overall benefit, an agent may reasonably and permissibly sometimes choose for herself to undergo that same harm to reap that same benefit.[18] It may be wrong for the surgeon to elect to break the hand of the unconscious patient to bestow the pure benefit but perfectly permissible and reasonable for a conscious patient to elect to undergo that same surgery. (A well-sighted patient might elect weeks of soreness and compromised vision for an operation promising heightened vision in the long run; or for aesthetic reasons, a patient might elect significant pain for a tattooed body.) If, ex hypothesi, the badness of the harm explains the first prohibition, why does the patient's consent make it permissible?[19] Why doesn't the badness of the harm instead render the consent subject to criticism?

Some explanation is provided by appeal to the patient's autonomy rights. Nonconsensual imposition of the harm would violate these rights, while acceding to her wishes respects their exercise. But without elaboration, this is not fully explanatory. It does not explain why the autonomy-right-derived prohibition applies only to imposing lesser harm to bestow greater benefits but not, prima facie, to imposing lesser harm to avert greater harm. Indeed, the case of lesser harm to avert greater harm has the inverse structure. Generally, where a lesser harm may be imposed to avert a large enough greater harm, this is sufficient reason to act even when a patient is unable to consent. But if the patient explicitly denies or withholds consent, for example, through a competently executed living will, then we may not act. The structure of autonomy rights seems already to presuppose an asymmetry between first- and third-person perspectives on harms and benefits, not to explain it.[20]

17. Also, suppose the patient really does appreciate the benefit conferred. She may even be, all things considered, glad it was bestowed. Still, if she complained that its nonconsensual bestowal also imposed harm on her and demanded compensation, she seems entitled to such compensation, whereas the patient in the Rescue Case cannot reasonably demand compensation. This suggests that those who aim to expose themselves to harm or to avoid harm-reduction efforts operate under a larger burden to make their intentions known.

18. Here I bracket issues about socially irresponsible uses of medical resources and assume they do not arise in this case.

19. Some doubt that it does. *See, e.g.*, Perry, *supra* note 5, at 1301.

20. An appeal to autonomy rights also fails to explain another closely related first-person/third-person asymmetry. As Scanlon observes, our assistance to others (whether morally mandated or not) may reasonably be targeted towards fulfilling their nutritional needs even if the recipients sacrifice their health to build a religious monument and would prefer

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

Moreover, an unelaborated appeal to autonomy rights reinforces an oversimplified account of autonomy rights. This view suggests a dichotomous picture: either the act that is impermissible without consent (here, causing harm to bestow a great benefit) would be an irrational one for the patient to select, or the act is the rational act to choose and the right protects a patient's prerogative to choose irrationally. This dichotomy places uncomfortable pressure on the defense of autonomy rights. At some point of substantiality, the defense of the right to engage in irrational, harmful conduct becomes awfully strained. Perhaps this defense can be made out, but it would be preferable to go easy on that strand of argument. Indeed, this simplified account eliminates ground that seems to exist. Namely, it may be wrong to impose harm on another because of the badness of harm yet permissible and reasonable for the potential recipient to opt for the harm. Hence it is a task for a theory of harm to address why consent should make such a difference, that is, why imposed harm may have a different moral significance from elected harm.

Thus a theory of harm should strive to explain why harm commands special attention. In particular, it should accommodate and (at least partly) explain (or explain away): the greater reason-producing and reason-requiring power of harm; the moral distinction between nonconsensual imposition of lesser harm to avoid greater harm and imposition of lesser harm to bestow greater benefits; and the related asymmetries between first-person election and third-person impositions of harm. As I argue below, standard accounts of harms and benefits cannot adequately accommodate and support an explanation of these asymmetries.

## III. STANDARD MODELS

The dominant picture of the structural relationship between harms and benefits portrays them as symmetrical. Typically, harms and benefits are conceptualized as occupying opposite ends of a sliding scale, like the positive and negative numbers, or as though they were mirror images of each other. Harms have a negative valence, and benefits have a positive valence. This symmetrical structure is usually unpacked by casting both harm and benefit as essentially comparative notions, invoking either a counterfactual comparison or an historical one. On the counterfactual view, one is harmed if and only if one is placed in a state in which one is worse off than one

us to support their mission. Thomas Scanlon, *Preference and Urgency*, 72 J. PHIL. 655 (1975), at 660. At least one of the example's lessons is that in my terminology, our welfare efforts may reasonably favor preventing and alleviating harm (i.e., nutritional deficiencies) over providing pure benefits. Still, recipients of such assistance might reasonably choose to weight their expenditures toward the fulfillment of their projects at the expense of their full nutrition. It cannot be argued here, as it might have been in the medical cases, that harm prevention is preferred because of ignorance of the recipients' wishes or because of limitations imposed by the recipients' autonomy rights, for the donors may target their assistance as they like.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

otherwise would have been; to be benefited is to be made better off than one otherwise would be. On the historical view, one is harmed if and only if one is made worse off than one was before,[21] and benefited if one is made better off than one had been. There is also, of course, the possibility of a disjunctive account in which one is harmed if one is made either worse off than one otherwise would have been *or* worse off than one was (with parallel moves made for benefits).

Such analyses require a metric of comparison, although many discussions of harm presuppose that there is some commonly understood metric without specifying one. I focus here on Joel Feinberg's interpretation, an interpretation that permeates much moral and legal theory. On this interpretation, harms and benefits are to be assessed relative to one's overall interests. One is harmed if one's overall interest has been set back, and benefited if it is advanced. Something counts as an interest if it is the sort of thing one has a stake in. That is, it is an interest if one's life is invested in it; its advancement makes one's overall life go better, and its decline makes one's life as a whole go worse.[22] This interpretation, coupled with the comparative framework, exerts a strong intuitive resonance. It is congenial to a dominant mode of economic-style thinking that is associated with cost-benefit analysis and has an attractive simplicity that may reflect some first-person reflections about how we make decisions and what we deem desirable or choice-worthy. Nevertheless, it cannot succeed as a plausible theory of harm.

The combination of the comparative component of these accounts and the interest-based metric of comparison creates two main problems: the accounts tend to both under- and overidentify harms and benefits, and they cannot comfortably accommodate moral asymmetries between harms and benefits.

## A. Standard Accounts and the Misrecognition of Harm

The standard accounts encounter a disparate set of difficulties in identifying harms and benefits. First, counterfactual accounts suffer predictable problems associated with examples of overdetermination. Just to rehearse a few familiar difficulties, suppose that if Jones had not shot Black, Smith would have done so or an avalanche would have felled Black. On either alternative, Black would be dead (or, for those who doubt that death makes one worse off, severely injured). Still, even though Jones's shot did not make Black worse off than she otherwise would have been, she suffered harm because

21. *See, e.g.*, THOMSON, *supra* note 6, at 261.

22. Feinberg is explicitly interested in legitimate principles of criminal liability and hence with the more narrow category of legal harm. Legal harm also requires as a component an element of wronging or wrongdoing that involves a rights violation. Feinberg, *supra* note 11, at 36. His conception of legal harm, however, rests on an underlying broader theory of morally significant harm.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

of it.[23] It will not help to resort to a modification that one is harmed if one is left in a worse state than one would have otherwise been in, assuming no comparable bad event would have occurred. For suppose that Jones shoots Black in the leg, but if he had not, Smith would have shot Black dead. On the counterfactual account, Black has not been harmed by Jones's shot. But it seems to me that Black's wounded leg and aching sores represent harms to her, although Smith's shot would have deposited her in far worse circumstances.[24]

In addition, the counterfactual approach produces odd consequences given certain forms of determinism. On some understandings of determinism, one must say that things could not have been otherwise than they actually turn out to be. If such a form of determinism were true, the counterfactual analysis of harm would have it that no one ever suffers harm for no one is ever worse off than she would otherwise have been. It seems quite implausible that the fact that harm does and has occurred could depend upon the truth or falsity of a variant of determinism. A theory of harm that is open to such a result is in some trouble.

Of course, these examples do not impugn a historical comparative account. In the cases just rehearsed, the victim is worse off than she had been. But such accounts have their own difficulties. Suppose I am drowning, Harry is about to save me, and you obstruct him. There I remain, drowning. I am not left in any worse a position than I was before, but still, I have been harmed by your action.[25] If efforts to remove me from harm are stymied,

23. *See also* Matthew Hanser's discussion of the relation between similar examples and the killing/letting die distinction. Matthew Hanser, *Why Are Killing and Letting Die Wrong?*, 24 PHIL. & PUB. AFF. 175 (1995), at 183–186.

24. Because the shot causes her to suffer harm, I would also say that Jones harmed her. If his act was performed with the proper motive, e.g., to avert greater harm, his act would not merit criticism. But it still inflicted harm. Had Jones performed it for a poor motive, e.g., the intent to inflict suffering, we would condemn him for harming Black, even though the alternative would have made her worse off. Some may claim that "causing to suffer harm" and "harming" come apart and that it is a conceptual truth that for A to harm B, A must worsen B's position. Thomson discusses a related case concerning historical accounts. THOMSON, *supra* note 6, at 263. She denies that you harm someone if you cause him (what would be a) lesser harm, e.g., amputate a leg, to eliminate greater harm (death or a more severe disability occasioned by a crushed leg). Her denial may be motivated by the permissibility of performing the amputation. But I do not think this excludes the possibility that the act harms. Even if harm were an essentially normative notion, the fact that it should be inflicted and that its infliction does not elicit charges of culpability or wrongdoing does not exhaust the forms of its significance. Thomson's position may also emanate from the tight connection she draws between suffering harm and having a *claim* against someone that they not bring it about, a connection I regard as too strong. *See supra* note 6. The amputated leg itself, prima facie, generates strong reasons to relieve the suffering associated with it, e.g., through pain relief and a prosthetic; further, it requires the same sorts of justificatory reasons associated with harm to bring it about, so I regard it as a harm even if one lacks a *claim* against the surgeon that she not amputate.

25. Why am I not worse off than before? Before I was about to be saved by Harry, and I am no longer in this anticipatory position. There are two difficulties with this maneuver. First, suppose I am left in the same position. That is, even after your intervention, Harry is still just about to help me. (And suppose you will continue to intervene to preclude his help from

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

I think it is fair to say that I am harmed by the obstructions even if I am not made any worse off than I was. Or suppose I am in pain, and you could easily alleviate my pain but you fail to. Even if your failure to assist me is not an act but an omission (and even if you are not morally responsible for my condition), the failure itself harms me even though it makes me no worse off than I was a moment before. Or you operate to remove a life-threatening blood clot, but the surgery results in a qualitatively identical blood clot. The new clot constitutes a harm to me, although substantively, with respect to my overall interests, I am not worse off than before.

Disjunctive accounts that combine counterfactual and historical tests are subject to this sort of counterexample: you operate to remove the blood clot, but the surgery results in a qualitatively identical blood clot. And if you had not done this, Harold would have. The life-threatening clot represents a harm to me, even though it does not make me worse off than I was or than I would otherwise be from the perspective of my overall interests.

Of course, there may be ways to refine comparative accounts to evade this or that difficulty posed by overdetermination. But I worry that such modifications will just dodge examples and patch holes in an ad hoc way without answering the underlying concern. These examples raise the question of why comparisons are salient. What seems basic to the moral significance of harm is the condition that one is in, not the condition one was or otherwise could have been in. One's broken limb will have a similar reason-giving and reason-requiring force whether or not it registers a comparative decline in one's condition. The concern is that even if modified comparative accounts can deliver the right answers in these cases, they will still be chasing after results. It will remain unclear why comparisons ought to be central to assessments of harm.[26]

This concern is underlined by comparative accounts' difficulties with acknowledging disabilities that appear at birth as harm. The historical

actualizing.) Your action does not eliminate Harry; you just preclude his putting his intentions into practice. It is unclear, then, in what way you have made me worse off than I was. Second, if we say that I am worse off because Harry was about to save me and be effective, then in essence the historical position has collapsed into the counterfactual account. For the account of why your action makes me worse off is not that I am in a worse position but, in essence, that I could have been better off but your action eclipsed this alternative. If interpreted that broadly, the historical account encounters the difficulties of the counterfactual account.

26. I do not suggest that comparative assessments are irrelevant for moral purposes. They are particularly relevant to choices between harm-inflicting or benefit-bestowing alternatives. If I am forced to harm one of two people and one person's harm is overdetermined, this gives me reason to harm her and to spare the one who may avoid harm entirely. But mere overdetermination will not necessarily excuse harm (or make a harmful activity nonharming). Other things being equal, it would be harmful and wrong to kill an innocent person even if she would be immediately killed by another were I to refrain. Further, I may have reason to provide a benefit even if others will provide it should I refrain. These reasons may include an obligation to contribute to a helping enterprise, to avoid free-riding on others' beneficence, or to establish or contribute to a particular relationship between myself and the beneficiary. *See also* Hanser, *Why Are Killing*, *supra* note 23; Matthew Hanser, *A Puzzle about Beneficence*, 58 ANALYSIS 159–166 (1998).

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

position has the strange consequence of holding that one born with a severely limiting disability is not harmed by that disability at all, although someone who acquires the disability just after birth may be greatly harmed by her condition. But what matters most in establishing what reasons we have to act is the position the person is in, not how it compares to his other life stages; the harm of the disability seems to constitute our salient reason to act.[27] Counterfactual accounts also mishandle a subset of these cases. They must contend that congenital, nontreatable disabilities that are significant enough to affect personal identity do not represent harm to the children who suffer them. Again, these conditions generate quite similar reasons to act that are close, if not identical in strength, to those that do not affect identity. Even if the comparative component makes some difference (although I severely doubt it)—so that there is some greater reason to treat children who become disabled after their identities are fixed than children with congenital, identity-fixing disabilities—the comparative component is not the main moral impetus behind relieving either group's conditions.[28] Treating the comparative component as central camouflages the greater number of aspects these cases have in common.

These accounts underidentify harm in two other respects. The sliding-scale, overall-interest approach to assessing harms and benefits encourages all-things-considered ultimate pronouncements that a person has been either benefited or harmed but not both. One has moved up the scale, down the scale, or stayed put. But this seems implausible. When a lesser harm is inflicted in order to avert a greater harm, one has still been harmed even if the net result is a positive one.[29] Or take a case in which an event both bestows upon one a tremendous amount of cash but simultaneously breaks one's leg. The breaking of the leg is nonetheless a harm, no matter how high up the welfare scale the financial windfall pushes one.

Second, most contemporary accounts of harm, including comparative accounts, exclude episodes of pain as harm unless the pain has larger, damaging by-products. Pain that also disables or distracts so as to impair pursuit or achievement of interests or projects counts as harm on these accounts. But simple pain itself has no comfortable place in these theories.

This is unfortunate. Consider cases in which I suffer short shocks of nonnegligible pain right as I am falling asleep, engaged in no activity, or even while I am diligently and effectively working. The pain causes no damage to my abilities nor does it distract me. I am able to continue working at the same degree of competence, but the pain unpleasantly commands attention, and I suffer. What I have in mind is not de minimis pain, like

27. *See* Shiffrin, *supra* note 14. Many discuss the problems associated with such cases. *See also* DEREK PARFIT, REASONS AND PERSONS (1984), at 351–442; Feinberg, *supra* note 11; Matthew Hanser, *Harming Future People*, 19 PHIL. & PUB. AFF. 47 (1990); Gregory Kavka, *The Paradox of Future Individuals*, 11 PHIL. & PUB. AFF. 93 (1982); McMahan, *Wrongful Life*, *supra* note 13; McMahan, *Cognitive Disability*, *supra* note 13; THOMSON, s*upra* note 6, at 262.

28. *See* Derek Parfit's discussion in PARFIT, *supra* note 27, at 366–371.

29. *See supra* note 24; s*ee also* Perry, *supra* note 5, at 1305.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

the slight physical annoyance involved in a minor hangnail—the sort that may naturally ebb in and out of one's consciousness without effort. Rather, I have in mind the sort of nonnegligible pain that commands attention and can be banished from one's consciousness only with effort. I think that sort of pain is a harm I undergo even if my other projects and the exercise of my abilities are unaffected by my suffering.

I have the straightforward intuition that such pains are harms but I recognize that this intuition is not universally shared. Independent of the intuition, it seems notable that there are the same sorts of moral reasons to avoid and alleviate pain as there are to avoid and alleviate other sorts of harm.[30] A theory of harm should be able to acknowledge this without committing to the more grand idea that the sufferer has a *stake* or *interest*, in the stressed senses, in not enduring pain. Avoiding nondebilitating pain may reasonably not be something that operates as a measure of how well my life goes.[31] Nonetheless, pain's imposition harms me.

Why do many believe that pain must have more detrimental effects on my abilities or my interests to be considered harm? For some, I suspect that the underlying motivation must be to avoid endorsing a view that recognizes all forms of desire frustration and nonfulfillment as harms. This is a laudable aim, but eliminating pain as a recognized harm represents too high a cost. It is intuitively implausible and it generates awkward efforts to accommodate the intuition that brute, nonnegligible pain provides moral reasons similar to those generated by other recognized harms. It seems worth exploring whether there is middle ground.

Finally, comparative accounts overidentify conditions as harm that do not merit the label. Both types of account will, under a suitable description of the states of affairs, identify as harm cases in which one merely loses or fails to receive a tremendous benefit. A billionaire's accidental loss of a thousand dollars will be said to be a harm to him, assuming he has a stake in his stockpile, as many billionaires do.[32] On counterfactual accounts, a missed

30. Both Thomson's and Raz's accounts seem awkward in this way. They acknowledge that pain gives rise to reasons that are like the moral reasons associated with harm, but yet they deny that pain is a form of harm. THOMSON, *supra* note 6, at 249–253, 264; JOSEPH RAZ, THE MORALITY OF FREEDOM (1986), at 413–414.

31. Ronald Dworkin argues that pains frustrate my experiential interests but they do not frustrate my critical interests and therefore do not make my life overall go better or worse. RONALD DWORKIN, LIFE'S DOMINION (1993), at 201.

32. One might modify the comparative account by appealing to thresholds in which only conditions that fall under a threshold of minimal overall interest satisfaction should count as harm. Such accounts may evade cases like the billionaire's monetary loss, but the price for evasion is too high. For it will then be difficult to acknowledge that the billionaire's broken leg is a harm to him should he remain above the minimal welfare threshold. One might then suggest that different thresholds cover different aspects of well-being. Ascents up the financial ladder do not compensate for physical harm.

Stephen Perry puts forward a more nuanced historical account that gestures in this direction, suggesting that historical worsening is a necessary but not sufficient condition of harm. Perry, *supra* note 5, at 1302. This would represent a step in the right direction, but without more elaborated details, it leaves important questions open and does not dispel the lingering

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

opportunity (not even a loss) to rake in another thousand will also count as a harm. This seems implausible and makes it puzzling why harm per se should attract any special moral notice. Of course, some opportunity costs are harms, but a better theory would register this fact without recognizing all opportunity costs as harms.

## B. Comparative Accounts and the Moral Priority of Harm

Not only do comparative interest-based theories under- and overrecognize harm, they lack the resources to accommodate comfortably the asymmetries between benefits and harms. Pure counterfactual accounts suffer an obvious and severe shortcoming. On these accounts, enduring harm and not being benefited, for example, suffering an opportunity cost, are identical. They are both conditions in which one is worse off than one would have otherwise been. And there is no room, likewise, to distinguish between the provision of benefits and aversions of harm. For with both one is better off than one could have been.

Historical accounts can draw some distinction between harm and the failure to be benefited. Harms will leave one worse off than one was before, whereas failures to be benefited may just leave one stationary. Still, the distance between the conditions associated with what counts as harm and those associated with the failure to be benefited may be extremely negligible. It is hard to fathom how such a negligible distance could account for the strength of our asymmetrical reactions and our sense of the special significance of harm.

Comparative accounts encounter another problem in making sense of the moral asymmetries between harms and benefits. Both accounts will count one person, Jones, as being harmed and another, Smith, as having been benefited even though both are deposited in the same situation and neither had a special claim to their starting places or their movements. Suppose Jones could have been or was at a higher status, $x+2$, and is brought down to x, whereas Smith could have been or was at a lower status, $x-2$, and is brought up to x. On comparative accounts, irrespective of the qualities of their resultant conditions, Jones will have been harmed and Smith benefited, although their conditions are identical. This problem becomes more evident if Jones moves from $x+2$ to $x+1$, whereas Smith moves from $x-4$ to $x-3$. On comparative accounts, Jones is harmed and Smith benefited, even though the former is much better off, all things considered.

Granted, in many cases, loss objectively does represent a harm, and sometimes a very significant one. But I suspect that when loss does represent a

suspicion that the significance criteria that infuse the necessary conditions will implicitly incorporate noncomparative considerations that do the fundamental work. Where to set the threshold reopens the question of what harm is that the standard account lacks the internal resources to answer.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

significant form of harm, this has more to do with the agent's legitimate expectations and investments in her position and their connection in turn to her projects and her identity than with the sheer difference between her current and former or potential states. Neither comparative account necessarily trades upon Jones's investments or expectations to conclude that Jones is harmed. If being deposited in a certain state can be either a harm or a benefit, depending on one's arbitrary prior or alternative positions, it again seems rather mysterious and implausible that harm itself matters especially.

These examples draw attention to the stance that such models, coupled with endorsement of harm's priority, must take toward norms of equality. It seems troubling that a theory of harm would suggest that achieving equality through redistribution of undeserved goods will both *necessarily* effect harm (merely by virtue of the losses or opportunity costs of the more fortunate) and that equality norms are thus prima facie in conflict with harm's priority.[33] Further, although achieving justice may sometimes involve tragic losses as we make transitions from nonideal circumstances, and perhaps some such losses are harms, comparative theories register harm here for the wrong sort of reason. They fixate on the comparative loss per se in finding harm rather than on the sort of loss involved; for example, they do not distinguish between the deprivation of an inflated savings account reserved solely for passing on to heirs, the loss of a third vacation home, and the loss of a treasured homestead that has served for generations as the focal point for family gatherings. All register as harm on the comparative account, but that assessment involves an implausibly high degree of tension with norms of justice.

These problems suggest that interest-based, comparative models of harm and benefit do not support and even represent obstacles to recognizing the asymmetries between harms and benefits. Are rival theories that use different metrics superior? Other comparative theories place stress in different places and aim to encompass wider or narrower ground. Judith Jarvis Thomson defends a comparative historical account that aims to include killings in the scope of harm (which she believes Feinberg's account is hard-pressed

33. This problem could be avoided by modifying the account to claim that one is harmed only if one is worse off than one deservedly was or would have been. This modification would, however, create a worse problem by eliminating the possibility of deserved harm. Michael Otsuka has suggested in conversation that the priority on harm might be understood as a priority on undeserved harm. Although this suggestion has some plausibility, there are problems with it. Even deserved harm is commonly thought to require special justification; witness the dissatisfaction with spare retributivist accounts and the desire to find further, positive reasons for punishment. *See, e.g.*, H.L.A. Hart, *Prolegomenon to the Principles of Punishment*, 60 PROC. ARISTOTELIAN SOC'Y 1 (1960); Alan Goldman, *The Paradox of Punishment*, 9 PHIL. & PUB. AFF. 42 (1979). Perhaps one's purposes need be less important if the harm is deserved or if the recipient's superior position is undeserved, but the fact that harm is involved elicits special justificatory burdens. The burden of justification for inflicting deserved harm or harmful deprivation of undeserved goods seems higher than that associated with deserved failures to benefit or failures to protect undeserved benefits.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

to do). Her account identifies harm with a worsening of condition that leaves one damaged, disfigured, or impaired.[34] Similarly, John Kleinig identifies harm as involving the impairment of welfare interests, where these are understood as the interests indispensable to the pursuit and fulfillment of characteristically human interests.[35] This emphasis upon damage or impairment, homing in on more central interests, protects against some of the problems that interest-based models encounter in explaining the special significance of harm. But not all of the difficulties are avoided.

Even when more focused upon central interests, comparative accounts still have difficulty explaining why the worsening, not the quality, of the person's condition is central. If the condition is what matters, it is unclear why any comparison to prior or possible alternative conditions is necessary.[36] Further, all comparative accounts have core difficulties explaining how harms could matter more than benefits and failures to be benefited. Finally, while focusing on the impairment of central interests marks an advance over wider setback accounts, central-interest accounts run the danger of excluding such things as pain and the actual frustration of projects (instead of just the deprivation of the abilities and general means for their pursuits) from the ambit of harm. More, then, needs to be said to delineate what counts as harm, to isolate why it matters, and to see if this explanation might also be made compatible with including such things as pain within our notion of harm.

### C. Appeals to Action

I suggest above that comparative accounts do not adequately distinguish between harms and failures to be benefited. One may object that while comparative accounts do not illuminate these asymmetries, they are not strictly incompatible with them. Harm's priority, one might suggest, has less to do with the conditions associated with harm and more to do with harming actions: agents' acts of harming are what have a greater moral significance than agents' failures to benefit. Often, the doing/allowing distinction is

34. THOMSON, REALM, *supra* note 6, at 227–271. She distinguishes pain and other forms of non-belief-mediated distress, such as nausea, but believes they may give rise to the same sorts of claims against others that harm does. By contrast, she argues that belief-mediated distress such as grief, indignation, disgust, anger, and disappointment do not themselves give rise to claims. I discuss this distinction in Section IV.D *infra*.

35. John Kleinig, *Crime and the Concept of Harm*, 15 AM. PHIL. Q. 27 (1978), at 31, 33.

36. Thomson may emphasize worsening because she, like many others, is interested in providing an account of harm as that which we have a claim against others' bringing about. *See supra* note 6. One might think, then, that worsening is not a primitive quality of the harmed state but is essential to causing harm in this sense. But if I fail to alleviate your suffering although I easily could, this could be plausibly described as a case in which my failure harms you and even I harm you although you are not made worse off; and further, you might have a claim against my doing this. But this issue lies outside my inquiry here. I mean to focus on the sorts of conditions that give rise to reasons to act and to refrain from acting, whether these conditions are brought about through human agency, its absence, or natural conditions, and whether or not these reasons are sufficiently strong to form the basis of a claim.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

represented, sometimes implicitly, as either embodying or explaining the distinction between harms and failures to benefit. Comparative accounts may render the conditions of harms and failures to be benefited indistinguishable or in uncomfortable proximity, but this does not entail that the acts (or omissions) of harming and failing to benefit are morally indistinguishable. The analysis of harm and benefit might not support an explanation for harm's priority, but complementary factors about agency and responsibility might fill the gap.

But, even assuming an important distinction between doing and allowing, it will not fully capture or explain the relevant asymmetries. For example, both failing to prevent harm and failing to benefit may be instances of allowing. Nonetheless, there is often a strong asymmetry between our reactions to them, even if the size of the benefit foregone is comparable to the size of the harm endured. In some cases, both may be permissible; still, other things being equal, failing to provide good nutrition or potable water to a person matters more than failing to provide him with a graduate education. Likewise, the harm/benefit asymmetry remains when one compares harming to the destruction or removal of a benefit, both examples of doing. The doing/allowing distinction does not explain or support either asymmetry.

One might rejoin that some doings are worse than others, and the same for allowings. Those doings that place a person closer to the negative end of the spectrum are worse than other doings. This permits comparisons of doings with doings and allowings with allowings. But coupled with the doing/allowing distinction, it will not clarify why, intuitively, both doing and allowing harm matter more than undoing a benefit *and* allowing the removal of a benefit. Breaking an arm and allowing an arm to be broken are, special obligations aside, morally more serious than interrupting the conveyance of a (comparably sized or perhaps greater) benefit, for example, the full funding of a person's project, and allowing the interruption. Both pairs involve both a doing and an allowing. Yet the moral division does not occur between only the doings and the allowings. An independent division arises between the harm-associated events and the benefit-associated events.

One could respond that the division occurs because the first doing/allowing pair is closer to the negative edge than is the second. But this rejoinder makes it evident that the doing/allowing distinction does little work to create or support the harm/benefit distinction. Further, the appeal to the negative edge is orthogonal to the senses of harm and benefit within comparative accounts. Failures to benefit, on the historical account, may be closer to the negative edge than harms (on the counterfactual account, they will be no different); failing to benefit a very poor person leaves him closer to the negative edge than actively taking a thousand dollars from a millionaire. Appeal to the negative edge will not, coupled with the comparative account, yield or explain a harm/benefit asymmetry.

Moreover, relying upon the relative distance from the negative edge to supplement the doing/allowing-based distinction bumps up against the case

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

above that involves placing a person very high on an overall welfare scale concomitant with inflicting a broken leg. Repairing or averting the break matters more than supplying (or protecting) a pure benefit, for example, a new stereo, a coveted bicycle, or a large amount of desired but unneeded cash to a less-well-off person. This holds even if the broken leg does not plunge the highflier down to a welfare level below the worse-off person. The harm has greater significance than the pure benefit, even though the harm-endurer is farther away from the negative extreme than the potential beneficiary.

## IV. AUTONOMY AND HARM: FURTHER DESIDERATA AND POTENTIAL DIRECTIONS

The problems with the comparative, interest-based models justify searching for another approach. The simplest alternative—a nonsymmetrical, list-based account that identifies harms with a set of evils (e.g., death, broken limbs, pain, disabilities) and benefits with a set of goods (e.g., material enhancement, sensual pleasure, certain sorts of goal fulfillment, nonessential knowledge, competitive advantage)—would differ structurally from comparative models; as such, it would be less prone to their difficulties. Without elaboration, however, it would lack the resources to explain what unifies them as harm, nor could it lend support to harm's priority and explain the asymmetries between harms and benefits.[37]

There are many reasonable starting points to begin thinking about a further-elaborated account. But the prominence of consent in the asymmetries suggests that the relation of the condition to the will of the agent who endures it may play a more integral role than comparative, interest-based accounts assign it. Further considerations about autonomy rights may yield additional reasons to credit this suspicion.

So far I have called attention to one aspect of harm and autonomy rights: that it is often impermissible for a third party to impose a harm to secure an overall benefit for a nonconsenting agent but permissible and even reasonable for that agent to make the same decision for herself. Another important feature of the relationship between harm and autonomy rights is the ability of an agent to choose courses of action that will not redound to her overall benefit but will redound to her overall harm. Although it is recognized

37. *See, e.g.*, Elizabeth Harman, *Can We Harm and Benefit in Creating?*, 18 PHIL. PERSP. 89 (2004), at 96–97; Harman, *Harming*, *supra* note 3, at 139. Harman, addressing what *actions* harm people, identifies harmful states as ones that are worse (in some respect) than life with a "healthy bodily state," i.e., states with cuts, burns, diseases, deformities, pain, and premature death. Although the characterization largely works for her purposes (except perhaps that it does not correctly capture what is harmful about rape), this approach does not explain why harm is limited to (physical?) health defects. Many harms we suffer do not necessarily affect our health. A more satisfying theory would accommodate this fact and explain the connection between states of compromised health and other sorts of harms.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

that avoiding harm is a major moral priority, most proponents of autonomy rights affirm that such self-regarding decisions should be protected and lie within the agent's purview. I call this stance "the broad view."

Although I concur with this conclusion, some of the standard arguments for the broad view are unconvincing or underelaborated. Popular conceptions of autonomy rights and what they protect seem ill-equipped to underwrite this guarantee. I believe, in part, that this is because challenges to autonomy rights emanate from two fronts, corresponding to two different but related ways that autonomy rights are deployed.

First, autonomy rights compose an important aspect of the pantheon of constraints on what may be done to particular individuals to promote aggregate welfare or, more generally, to promote the interests of other people and the community. The corresponding challenges ask "why" questions, such as why a person's free choices or her bodily dignity or her privacy should be respected even if many innocents will suffer as a consequence (say, of an ill-considered but persuasive political speech or an unwise vote in an election or a selfish refusal to submit to a timely search without a warrant). Second, they act to insure an individual's unimpeded ability to make decisions regarding herself even if her efforts threaten her welfare and perhaps even her autonomous capacities. This position confronts incredulous challenges about why an individual's decisions that will harm herself—even in ways that she admits are harmful to her—should be respected.

Often the orientation of defenses of autonomy rights gives the impression that challenges on the first front are being answered. And then, without explicit recognition, this sort of defense is uncritically relied upon to do the work against challenges on the second front. But the second front poses some distinctive challenges that may not be entirely met by defenses of autonomy rights that stress, however correctly, the inviolability of the individual against the value of the whole. This failure to attend to these distinctive issues may account somewhat for the persistence of skepticism about the range of the individual's right of self-determination even among otherwise staunch nonconsequentialists.

After elaborating on this problem, I suggest below that autonomy rights and harm avoidance should stand in a more cooperative, integrated relationship than many accounts represent. This may both help to defend the broad view and suggest the direction we should take to arrive at a plausible theory of harm.

## A. The Difficulty with Character-Based Accounts

A common articulation of the broad view is to liken autonomy rights to trump cards. Structurally, as rights, they simply overcome the general directive or goal to act to avoid harm to agents. This metaphor delivers a great

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

deal of satisfaction. One plays the trump card, and there is no further appeal. In a way this seems fitting—the image reflects our conviction that our lives are ours to squander as we please and not to be tampered with—even for our own good.

But the strengths of the trumping image are also its weaknesses. It is an articulation of the consequence of a certain commitment to autonomy. It is not really a defense of that view. It offers little explanation to the skeptical, and its sheer force may seem like bravado to the unconvinced. It leaves unexplained why the autonomy right, with this sort of trumping power, has the scope to extend to elect harm and why the significance of harm does not itself ground a right against being harmed that has its own trumping power, even as against an agent's own consent.

That is, the trumping image suggests a kind of orthogonal opposition between autonomy rights and harm avoidance. The two values run on opposite tracks, in some cases they collide, and when they do, the autonomy right has right of way. This model places a lot of pressure on the defense of autonomy rights. Of course, fleshing out the account of autonomy's importance can go a long way to vindicate this claim. But it is a tall order to show that autonomy rights, understood as distinct and separate from harm avoidance, are so vital that they will always necessarily outweigh or trump the importance of the harm that an agent will suffer as a consequence. If autonomy and harm avoidance are distinct enterprises, and harm is sufficiently important to ground key moral principles, it would be surprising that even when a great deal of harm is at stake and the decision in question engages autonomy values only to a minor degree, still the lexical priority of autonomy clearly holds. Why should the one central moral value typically or reliably trump its distinct yet vitally important counterpart?

As long as autonomy rights and freedom from harm are conceived of as disparate, even incommensurable items, this question will arise. If autonomy rights and harm avoidance were understood as related, however, the justifications for permitting agents to elect harm might be easier to make out.

There is, of course, a familiar move to connect autonomy to harm avoidance. Proponents of autonomy rights may assert that leading an autonomous life is a core interest of agents. Although a particular decision made by an agent may in the short run appear not to further her overall interest, in the long run it will, because it will contribute to her higher-order interest of leading an autonomous life.

Although I subscribe to the basic spirit of this move, I worry about its power for two reasons. First, it is not obvious that it answers the justificatory worry simply to declare that autonomy represents a higher-order interest. Second, if autonomy is seen as an interest of the agent, it raises the same sort of worries—why is it not possible that on occasions the agent's other interests will surpass it in importance or also lay claim to being higher-order interests?

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

There may be a different way to bridge the connection between autonomy rights and harm avoidance. Above I suggest respects in which the typical conception of harm is too distant from its normative significance to vindicate it. A converse aspect of the problem may be that the cited justificatory foundations of autonomy rights have been too tightly tied to an ideal of the person associated with their wise exercise. Reconceptualizing both harm and the justificatory underpinnings of autonomy rights may reveal greater connections between them.

Let me articulate the problem in terms more specific to prominent accounts of autonomy's value, representing that value as protecting an agent's ability to reflect on, develop, and express a distinctive character and set of life-characterizing ends and values. Many such accounts suggest that to have autonomy is to have the capacity to reflect critically upon one's desires (and to have this critical reflection effectively influence one's desires). Strangely, this account of autonomy does not make central the *implementation* of one's world-regarding desires. That is, these accounts emphasize that the successful implementation of one's *second-order* evaluations of one's desires may be part of one's autonomous capacities. But they omit from the key features of autonomy the ability and opportunity to implement one's first-order desires—to put them into action or to make them one's intention in action.

This omission makes the picture of the autonomous agent strange and the fit between autonomy and autonomy rights, as typically conceived, rather mysterious. For it seems that on this view, an agent who was able to align her desires and intentions with her critical evaluations would be autonomous even if her every move was impeded by others and she was unable to implement her desires. True, her capacity to reflect on her alternatives and to regulate her character makes her the sort of agent that *merits* autonomy rights. Still, if she were fully impeded in this way, she would not live a fully autonomous life or fully realize the values of autonomy. Further, given this conception, the core autonomy rights should be comprised dominantly of rights against mental interference, for example, rights against deception, fraud, brainwashing, and manipulation—that is, rights against mind control and other actions that would impede the free, authentic process of self-evaluation and self-determination of character. If autonomy does not necessarily involve the ability or opportunity to implement non-character-regulating desires, then this conception would not give illuminating support to the protection of choices generally associated with autonomy rights, for example, to decide on certain courses of action, to accept or refuse treatment, to act out of character, and even to engage in imprudent *action*.

An obvious response is to supplement the character-based account and add that expression of character, and not only its reflective, authentic formation, is a central aspect of autonomy or its value. Adding this component makes a difference, but it does not entirely solve the problem of imprudent action, because that fix is not responsive to the problems that arise from primary, daily exercises of autonomy rights that are not directly connected

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

to character expression. Character-based accounts of autonomy, even if they assign a significant place to the expression of character in action, become awkward when they confront "small" cases—the daily, picayune decisions that autonomy rights typically protect. Consider the decision to smoke another cigarette, to eat one's sixth chocolate bar of the week, to drink another glass of wine despite some work obligations for the evening, or to order French fries instead of salad. In some respects, these sorts of decisions represent the most frequent exercises of autonomy rights we encounter, ones that we hold very dear.

Of course, bans on intimate unions, career decisions, and political speeches are more injurious and egregious. But interferences with these small but more frequent decisions rankle and can rankle quite severely. And their rankling echoes the insult we feel when larger decisions are taken out of our own hands, whether for our own putative good, for the benefit of others, or for no reason at all. But character-based accounts have a difficult time acknowledging these cases as the bread and butter of autonomy accounts. They do not obviously involve the selection or expression of a character or self-defining values.[38]

One may respond on behalf of character-based views that the development and exercise of character may not be an entirely conscious, intentional, scripted effort. The development of habits over time through a series of small incidents and episodes is as central to character development and expression as more momentous, intentional affirmations of values or large-scale decisions. The repeated failure to resist the extra dessert may be the process through which one fails to develop continence and embarks upon incontinence or even gluttony—whether this is one's intention or even one's self-conception. Small decisions are the sites at which character development occurs and is expressed.

There is much truth to this reply. Nonetheless, this rejoinder does not explain why autonomy rights, underwritten by the character-based explanation, cover decisions that are both isolated and out of character—cases in which the decisions do not represent the beginnings or even the possibility of character change or evolution, such as nonrecurring decisions or choices at the end of life. It could be said these choices speak to the aspect of autonomy that involves the expression of character, for the decision may represent a test of constancy, or one might say that although the decision to eat an extra chocolate is not itself an expression of character, the protected option to indulge provides the *opportunity* for the meaningful development, invocation, and exhibition of distinctive character traits and virtues, for example, of restraint, modesty, or even neurotic fastidiousness. To interfere with such episodes would deprive agents of these opportunities and so would interfere indirectly with autonomy.

38. I discuss related issues in Seana Valentine Shiffrin, *Autonomy, Beneficence, and the Permanently Demented*, *in* DWORKIN AND HIS CRITICS 195 (Justine Burley ed., 2004).

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

These things may indeed be said, but must the defense of choosing to have another Snickers bar, whether prudent or not, really rest on viewing the choice as an ordeal of character? It is misleading to say that the decision is protected by an autonomy right because one option represents a valuable expression of autonomy and the other option makes the former meaningful. It is not just the abstemious decision that represents an exercise of what we mean to protect through autonomy rights. The out-of-character decision to overindulge is also a direct exercise of these rights. Character-based accounts seem ill-equipped to acknowledge this or do so in a convoluted manner.

Further notice that the right holds and seems to have a similar character in contexts of small choices that do not represent, even in high metaphorical mode, the opportunity for character-based missteps or revelations. Suppose I have decided to have the chocolate bar (and suppose it is not imprudent) and you interfere with my choice between the Mars and the Milky Way because you think one is clearly superior. This would be an objectionable albeit minor intrusion. We understand the daily exercise of autonomy rights as involving such mundane, unencumbered choices as those selections of Hershey's Dark or Milk, white or black pants, long hair or short, the College Avenue or the Piedmont Avenue route. Of course, sometimes these decisions encode personal and cultural symbolism, but again, that seems too indirect and too intellectualized a justification for the protection of such choices.

Thus I find character-based accounts unsatisfyingly incomplete. The values they emphasize do, of course, matter. Developing and expressing a distinctive character and a moral personality may be the most important ends toward which autonomy rights are directed and the instances in which they are strongest. They may do substantial work in grounding the right, explaining its contours,[39] and defending autonomy rights on the first front, regarding claims made on the individual by considerations of social welfare or others' needs. But decisions that involve character development and expression do not seem like a different species from the smaller decisions I mention above. Interferences with both strike us as offensive intrusions on our person. An account of autonomy rights should be able to represent these cases as on a spectrum without convolution.

An account of autonomy's value that underwrites autonomy rights, so understood, must give more weight to the significance of the power and opportunity to *implement* one's first-order, world-regarding desires, intentions, plans, and projects, however substantive or trivial. An account that gives central prominence to the ability to exert control over the contents of one's

39. They may be most useful in explaining where and why autonomy rights generate strong positive duties. So we may have duties to contribute to creating conditions in which people may develop their capacities, by providing adequate options, sufficient educational training, and a tolerant atmosphere, etc. These duties may be linked to the character-based aspects of autonomy's value.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

experience as well as some sphere of one's environment seems better able to span the range of behaviors protected by autonomy rights.

A broader account of autonomy's value would locate it in the sheer exercise of agency—in an agent's efforts to exert her will to influence, adapt, manipulate, and conform her experience and environment to fit her will. The value of autonomous action lies both in the object of achieving harmony between an agent's will and the contents of her experience and environment, and in the means of its achievement—that this object is effectuated through the exertion of the agent's will. At their core, what autonomy rights protect is an agent's ability to exert control over her mind, her body, and some aspect of her environment. The central idea is one of an individual's having a domain that is understood to be hers and subject to her unique, exclusive authority. This permits her the opportunity to construct (or heavily contribute to) the environment she inhabits so that it bears her distinctive mark, is flush with her underlying will, and is the product of her authorship. The construction and expression of character and intellectual virtues are of extraordinary significance on this account, but they do not exhaust its motivations.

Its roots lie foremost in a sense of the significance of the separateness of persons and the value of their separateness. Autonomy rights respect that an individual's will is distinct and separate from others' by respecting a domain in which that will is sovereign. They respect that the individual has a special, intimate relation to her mind, body, experience, and environment that she must especially endure, rendering it fitting that she and not others (whose relation and exposure to these experiences and conditions is more distant and indirect) exerts control over it—whether that control represents an especially substantive expression of character, rationality, or other intellectual virtues.

I suspect that an account along these lines, coupled with a suitably structured theory of harm that is more responsive to the previously articulated desiderata, would dovetail and facilitate a better explanation of why agents' autonomy rights extend to encompassing some forms of self-elected harm that could not be permissibly imposed by others.

## B. Possible Directions for a Theory of Harm

Unfortunately, I lack a fully satisfying account that both makes better sense of why autonomy rights extend to protect agents' deliberate self-exposures to harm and explains the other first-person–third-person asymmetries involving the imposition of harm. Instead, I have a tentative hypothesis about a direction in which to proceed.[40] I suspect that what we classify as harms

40. A related account of harms and benefits, but one with more developed and pronounced Aristotelian features (and one that takes a historical comparative form), is offered in David Velleman's fascinating and thoughtful three-part article on future people and procreation.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

must be the sorts of things that could reliably give rise to reasons to avoid, prevent, or alleviate them consistent with our maintaining the social and interpersonal conditions of meaningful autonomy. This in turn suggests that what a harm is will bear a strong connection to the underlying individual's conditions of autonomy and whether the state she is in constitutes an obstruction or impediment of her control over substantial elements of her experience.

One approach that might achieve these aims would involve a will-oriented conception of harm on which harm involves a distinctive sort of frustration or impediment of the will or of the ability to exert and effect one's will. An advantage of this hypothesis is that it may vindicate the intuition that pain amounts to a form of harm, without appealing solely or brutely to its unpleasantness. Consider unpleasant pain that does not disable or distract one from pursuing fully one's endeavors but is clearly felt and suffered. What seems most objectionable about this sort of pain is not merely that it is unpleasant or one does not desire it; rather it is that it forcibly intrudes and occupies an aspect of one's consciousness, is insistently present, and imposes an experience that is alien and directly at odds with one's will. Absent other resources and efforts, one must just endure or bear the noxious experience—in this way, one is passive towards it. Being trapped within a close, present experience to which one is passive and resistant instead of being actively engaged and identified with, or at least comfortably indifferent to, seems central to its objectionableness and, I would hazard, to its harmfulness. Unchosen pain is altogether intrusive and forced. The case of chosen pain is, fittingly, more mixed. Electing pain renders the cleavage between the will and the experience less severe, because the pain nests within some larger purpose or endeavor with which the will is engaged; still, the nature of felt, unpleasant pain nonetheless necessitates some nonnegligible estrangement between the will and a prominent aspect of one's experience. Usually, then, chosen pain represents a case of self-elected harm.

This description corresponds to many pain-management strategies. Many sufferers attempt to withdraw into their thoughts and to avert attention from their experience to avoid or dull the overt sensory aspects of endurance. Others scale back their activities to avoid the pain, or enact psychological strategies of resilience, some of which require emotional steeling and distancing from one's experience. Others attempt to rationalize the pain. They try to understand its sources and even to justify its presence. This may represent an effort to connect to this foreign presence and integrate it into a system of explanations, rendering it less alien and partly worked into the content of one's life and one's rational activities.

On this account, pain's unpleasantness plays a key role, but is not the full and ultimate explanation of pain's harmfulness. Insistent, involuntary, and

See J. David Velleman, *Persons in Prospect: Parts 1, II, and III*, 36 PHILOSOPHY & PUBLIC AFFAIRS 221–288 (2008).

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

ineradicable sensory pleasures may also have these objectionable qualities. Their pleasantness makes the experience intrinsically easier to integrate with one's will; whereas pain's unpleasantness helps to explain why pain reliably creates cleavage. But their pleasantness does not guarantee this. Unwanted, involuntary pleasures can remain foreign invasions just to be borne. Think of protracted bouts of tickling, unwanted giddiness, or a pleasant but persistent, insistently present and unwanted snippet of music (whether from your neighbor's loud stereo or from a musical recollection that cannot be dismissed from your consciousness). Consider also a strong sensory pleasure that results from an activity one regards as immoral or base.[41]

May a similar characterization be given of other sorts of harms? Perhaps. In many cases the relevant estrangement will not (only or necessarily) involve a *subjective* clash between one's will and one's conscious experience, as it does with pain. There are analogous objective components of harm. Some harms involve a significant chasm, conflict, or other form of significant disconnect between one's will and one's life. I understand one's life here to range beyond conscious awareness to include such things as one's relationships, projects, and the aims of one's efforts.[42]

Such a characterization may provide a way to understand what is so seriously objectionable, for example, about the substantial frustration or destruction of an agent's projects even when she remains unaware of it. Although it is not the principal reason we do or should engage in them from the first-personal point of view, an agent's projects play a substantial role in making her life her own and in providing the forum in which she actively surmounts or overcomes passive, given, burdensome aspects of existence that are fixed, imposed. That is, an essential aspect of our existence is that we are embedded in a range of unchosen conditions, including physical pain and physical limitations, many unchosen personal and moral relationships, many unchosen moral obligations, and so on. One's serious, large endeavors in effect help to render one's life distinctly one's own—to select its important themes and to make it an arena that is chosen, affirmed, one's own home.

In other words, a collateral side effect of one's dedication and immersion in one's projects (and relationships) is partly to remove oneself from (or

41. As Greg Keating observed to me, in some contexts, sexual and narcotics experiences may bring pleasure but yet be overwhelming in ways that shade or threaten to teeter into this characterization. In the good cases, the absence of control and the totalizing nature of the experience are sought and well realized; the experience is more liberating than invasive. In bad cases, the borderline is crossed, and though pleasurable in some respects, the experience becomes alienating and sometimes frighteningly so. And, of course, addiction and other sorts of dependence may be risked that involve a more thorough compromise of one's agency and autonomous capacities.

42. The claim that the contents of one's life exceed the boundaries of one's beliefs and experience has been widely discussed in the literature on death. *See, e.g.*, THOMAS NAGEL, *Death*, *in* MORTAL QUESTIONS 1 (1979).

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

dampen the prominence of) harmful states. Hence their unraveling does harm to a person not because it disrupts the fulfillment of important reasons and strong desires but because it impedes a significant aspect of agency that in part removes us from states of endurance. Complete success is not necessary for this aim's achievement, but the destruction of a project or the substantial interference by another may undermine this process; it may render an aim hopeless or futile and thus make a life structured around it importantly foreign.[43] The destruction or ruin of an agent's serious projects substantially clashes with what she has willed and invested herself in; the will may be seriously frustrated, even when it and conscious experience are in harmony, as when one falsely believes one's projects to be successful.

Being a victim of fraud provides another example. One may be unaware of the fraud, but the manipulation involves being treated in a way that is sharply inconsistent with what a rational agent could reasonably will, for it is inconsistent with respect for one's status as an independent, autonomous agent and one's efforts to exercise one's will accordingly. When the manipulation succeeds, one's life may be shaped by decisions to which one has not been an authentic party.

Similarly, others' lesser interferences with one's projects may also represent harm to the extent that they render one subject to the will of another. This signals a respect in which rights violations may represent harms, even when the condition that results would not, absent the context and cause, represent a harm. (A virtual stranger takes it upon himself to borrow a book from your library without your permission for a day. You had resolved to read it for sentimental reasons to mark a minor anniversary. The stranger stymies your plans. Given the property rights violation, we might say his theft represents a harm, even if otherwise the frustration of this minor plan, say by a distraction or a swelled library door that would not budge, would not.)

43. The frustration or failure of certain projects that essentially involve risk or competition raises further complications that I can only briefly describe. When such projects fail or are frustrated because the risk is realized or the competition is lost, we may have reason to deny that failure or frustration so caused represents a harm. Or in the alternative, we might regard their failure or frustration as self-inflicted harm. I am tempted to say that properly understood, such projects really involve *attempting* to surmount a risk or *attempting* to best others. Thus the failure to win, for example, therefore does not involve a failure or impediment to mounting the attempt. One might also say, perhaps, that the failure of that sort of project for that reason does not deposit one in a condition foreign to one's will; where the essential value of the project involves confronting risk or competition, the realization of the downside may be disappointing but is an aspect of the world one embraces, as opposed to the sort of destruction caused by a hurricane, earthquake, accidental plane crash, or malicious destruction by another. Thus I dispute that, as Scott Hershovitz seems to maintain, the intuition that a loss or setback caused by competition is not harm may only be explained by embracing the idea that harms are necessarily the product of wrongs. Hershovitz, *supra* note 3, at 1166–1167. A classic discussion of the relationship between harm and losses sustained through competition appears in JOHN STUART MILL, *On Liberty*, *in* ON LIBERTY AND OTHER ESSAYS 104 (John Gray ed., Oxford University Press, 1988) (1859).

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

These are, admittedly, hard cases. I am inclined to say that the rights whose violations we tend to regard as harms are either ultimately harm-based rights or autonomy-based rights, that is, they are rights directly against being harmed or rights that protect one's prerogative to exercise control and supremacy over a specific sphere of one's life. If one's *autonomy* right is violated but one's resultant condition is not in itself a state of harm, the violation itself may still be a harm, because another's will is being imposed upon one in an arena in which one's will is supposed to be supreme.[44] This embroils one in a form of passivity, for one is subject to the will of another and must endure another's choice in a domain that is supposed to involve one's own direction. (By contrast, your violation of my institutionally granted and institutionally oriented right, aimed to ensure that each of us has the ability to play a role in a scheme of checks and balances for the overall social good, might constitute an affront but not a harm to me.)

A like-minded explanation might be explored with respect to why disabilities, injuries, and illnesses are typically harms. They usually involve the subjective aspect of harm because they forcibly impose experiential states that clash with one's will, creating the same sort of chasm between one's will and one's close, lived experience that surfaces in the case of pain. They also typically involve the objective aspect of harm. Such conditions often significantly impede one's capacity to achieve substantial congruity between one's will and one's life. Unelected death also seriously interferes with agency.[45] By constraining the duration and the contents of one's life, it forces a particular end to a person—making her passive with respect to that central aspect of her life as well as to the resolution of the projects and relationships it undermines and curtails.[46]

44. If we conceive of promises as transfers of power or decision-making authority, then we might be able to conceive of breach of promise (and breach of contract) as a *harm* without having to resort to a counterfactual comparative theory. For worries as to whether noncounterfactual theories can accommodate breach of contract, *see* Goldberg, *supra* note 5, at 1325.

45. Matthew Hanser suggests that a state-based approach cannot accommodate the idea that death is a harm because the person who has died no longer exists and so is not in a bad state per se. Hanser, *Metaphysics of Harm*, *supra* note 3, at 437–440. I disagree. First, being subject to death can represent a harm. One's mortality coincides with one's existence, even if one's death does not. That reply, however, does not capture the sense that not merely being subject to death but actually the fact of death is a harm to one. But if we think the boundaries of a person or her life may extend beyond her corporeal existence, then she may suffer harm even when she no longer exists. That is, Hanser's argument seems to presuppose an existence requirement that is in conflict with other familiar ideas, such as that some events after a person's death may harm them by frustrating their core projects. Importantly, the way death harms one is, on this account, unsurprisingly an extreme version of the way other harms operate; the possibility of exerting and implementing one's will is not merely impeded or obstructed, but eliminated. One might also take a different route and argue, as Judith Jarvis Thomson does, that death is reasonably thought of as a special case. Thomson, *More on the Metaphysics*, *supra* note 3, at 453–456. *See also* Matthew Hanser, *Still More*, *supra* note 3, at 462–465.

46. So this approach is not precluded from counting as harms natural deaths, ordinary inabilities, ordinary pains, naturally contracted diseases, or the susceptibility to them because of their naturalness or ordinariness. They impose states that generate strong reasons to act to avert or relieve them; their strength is not lessened just because they happen to everyone or are predictable. Still, the context may matter enormously as to which states function as harms;

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

More controversially, it may be claimed that other harms will involve an analogous, significant chasm between one's will and critical aspects of one's circumstances: I have in mind, for example, serious ignorance and its ramifications. Severe ignorance and misperception can count as forms of harm, I think, because they mean that one's will is not engaged with the actual, central facts about one's circumstances. In cases of false consciousness, for example, there may be no conflict between one's will and one's life and experience—one may get what one wants as one conceives it and know that one has gotten what one wants as one conceives it. But severe misunderstandings can create a significant disconnection between one's will and the life one leads, on the one hand, and the nature of the circumstances that affect one and that one creates, on the other; moreover, one's will may be altered or made inauthentic by one's ignorance—a person's minimal understanding of his situation might substantially alter his conception of what he chooses. In these ways, serious ignorance will make one passively subject to one's circumstances; one will be situated within a milieu and subject to a range of forces and influences that one does not properly understand, much less react to in a deliberate way.

All of these conditions register, I think, as harm. Although they differ from one another, they might be thought to share in common the feature that they place agents in a relation of conflict with or estrangement or alienation from significant aspects of themselves, their conscious experience, their lives, or their circumstances.[47] This makes them, with respect to these conditions, more like involuntarily passive endurers of their lives than agents who are active and engaged or at least comfortable and familiar

an atypical condition in a particular social context may operate as an inability, whereas in another social context differently designed, it might not; being deaf in a dominantly hearing community is often a harm, whereas being deaf in a predominantly deaf community or in nineteenth-century Martha's Vineyard (in which sign language was widely used by the hearing and the deaf alike) might not be (or might represent a significantly lesser harm, depending on whether one is inclined to regard a lack of access to music as a deprivation of a benefit or as significant form of estrangement from the world and therefore a harm). *See* NORA ELLEN GROCE, EVERYONE HERE SPOKE SIGN LANGUAGE: HEREDITARY DEAFNESS ON MARTHA'S VINEYARD (1985). Although I regard it as a strength of the approach that it can recognize quotidian conditions such as ordinary pains and diseases as harms, others see this as a liability and worry that this approach overidentifies harms. *See* Frances Kamm, *Baselines and Compensation*, 40 SAN DIEGO L. REV. 1367 (2003), at 1384–1385; I. Glenn Cohen, *Intentional Diminishment, The Non-Identity Problem, and Legal Liability*, 60 HASTINGS L.J. 347 (2008), at 373–374.

47. In this way, a will-centered account may capture some of what is attractive about the standard, interest-based model. As I am arguing, interest-based accounts seem overbroad and threaten to veer too close to desire-fulfillment accounts that pose a challenge to a maintaining harm's priority. But still they exert a persistent pull. Part of their attraction, I think, is their connection, through the emphasis on what the agent has a stake in, to the agent's will. An approach that locates some essential connection between harm and the will may accommodate some of what drew us toward standard accounts in the first place. For example, it has the resources to acknowledge that certain severe losses are harms without claiming that all losses are harms. Some severe losses of position or comparative setbacks to interest will not land one in a position that is intrinsically harmful but may represent a harm for that individual because the loss will place her in a circumstance that is foreign to her, given her self-conception, expectations, and investments in her present situation.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

with the main and prominent components of their lives and experiences. The most serious harms seem to involve states or conditions that obstruct, frustrate, or directly conflict with an agent's cognizant interaction with her circumstances and her ability to lead a life within those circumstances that is distinctively and authentically hers—a life that is more than merely that which is watched, marked, endured, or undergone.

This sense of passivity that is connected to passive endurance differs, then, from a more familiar notion of passivity. The familiar notion is one that contrasts with bodily activity, mobility, or control exerted through agency. But one may be passive in my sense while writhing in pain or while acting violently under the grip of a delusion or a set of uncontrollable passions. And in the sense that I am after, one is not necessarily passive when one receives assistance or watches television. Whether one is passive depends upon one's will and its relation to one's state. One can sometimes avoid passivity even in the face of a lack of control by claiming, accepting, or endorsing (for reasons) what one is subject to—that is, by being mentally engaged with it, authentically accepting it, or regarding it as a welcome aspect of one's life. Passivity, in the sense I am after, has to do with conditions or events merely happening to one in the starkest sense; they just wash over or operate on one without one's being engaged or identified with them or actively responsible for their occurrence. One is in a sense just a bystander (sometimes even unwitting) to an aspect of one's own life.

This identification of harm with conditions that engender sufficient disconnection with the will does not rely on a related but more robust conception of the good life as involving maximal development and exercise of one's powers. To avoid harm, one need not manifest as much self-authorship, self-reliance, and independence as possible. The claim is not that all unchosen conditions or experiences are harms. Many or most such conditions are not sufficiently central or prominent, and many are susceptible to endorsement, connection, or just simple compatibility with one's will. Receiving gifts and having one's food prepared and delivered do not involve passivity in this sense. To occasion the sort of passivity I have in mind, the condition must be sufficiently prominent or important and it must be sufficiently antithetical or foreign to one to occasion either a substantial clash or a severe gulf between it and one's will.

Nonetheless, there is a close connection between this sense of passivity and the more familiar sense that contrasts with activity. One can be passive, in the sense of not being active, to many local conditions without being harmed. But generally, avoiding harm and passivity in my sense throughout one's life does in fact require a great deal of activity and exercise of one's autonomous capacities. To have a life that one is comfortable in, that feels like one's own, and that one may identify with, it is likely that in many arenas one has to be active in the more robust sense. One has to exercise autonomy in large and small ways, developing and deploying one's capacities for choice and implementation. Having an active hand and exerting actual control are

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

often required for one to feel that one's states are specially connected to oneself—that they are one's own in a proprietary sense instead of merely being something one is deposited in or that just happens to (against?) one. It is no accident, then, that autonomy is closely connected to harm avoidance and that many impediments to autonomy, such as injuries and disabilities, represent forms of harm.[48]

On an account like this, the justificatory foundations for autonomy rights bear a close relationship to the explanation for harm's priority. Both emphasize the significance of an agent's achieving some degree of harmony between the contents of her life and her will. Indeed, I think it is an important aspect of the justification of strong autonomy rights that they afford individuals the opportunity to avoid a wide range of harms. But although autonomy rights and the means to avoid harm bear a close relationship, they are not identical. It is possible for an individual to exercise her autonomy rights so as to do herself harm. Further, the range of one's autonomy rights may not be as extensive as the range in which one may suffer harm, on the one hand, and in many ways extend further than the range in which one may suffer harm; they include the ability to select and pursue benefits as well as to pursue moral and other-regarding ends.

## C. Advantages of a Will-Oriented Account

Adopting a more will-oriented account of harm may help to further our understanding of harm's special significance and its priority. Rendering (or threatening to render) one passive to or alien from a close, mentally or physically intimate experience or to the larger contents and structure of one's life seems especially significant. That which constitutes or substantially contributes to a person's being trapped in and alienated from that which she must also undergo seems particularly troublesome. It is, I think, clearer why such conditions would matter more than mere setbacks of interest and why they would matter independent of one's prior or alternative position.

A view of this kind may make sense of harm's special status in moral theory. It also holds out the promise of contributing to an explanation of

48. Thus I am sympathetic to Joseph Raz's account of harms as impediments to autonomy. Raz, *supra* note 30, at 369, 448. But there are reasons to be cautious about this characterization. It can be understood in a way that is overbroad. Many but not all obstacles to self-authorship are harms. Impediments to full self-authorship may well be undesirable, and their removal will often represent a benefit. But if one's relation to one's experience, life, and circumstances is sufficiently cognizant and integrated, then it may be desirable (for some people) to be able to replace some of unchosen elements, but it might not be essential for that agent's avoiding harm. If every obstacle to complete self-determination were considered a harm, then the account would face renewed problems about explaining harm's special priority. Second, the view can be underinclusive. By focusing solely on the abilities and obstacles to developing and exercising one's autonomy, it tends to exclude noxious, intrusive, but temporary conditions such as pain.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

the asymmetries between benefits and harms, assuming we also refine our idea of benefits along the lines I suggest above.

Conceiving of harms and benefits this way enables us to account better for the asymmetries between harms, benefits, and failures to be benefited and thereby to avoid the structural pitfalls of the comparative accounts. If the account of harm itself recognizes qualitative differences between harms, pure benefits, and failures to benefit, this makes it more readily conceivable that asymmetries between them could arise. More substantively, it seems clearer why harms matter more than failures to benefit. Benefits are of course desirable and choice-worthy, but without them, one is not left in a position akin to or in some other way as morally serious as a condition of harm. This alone goes a long way toward explaining the first asymmetry, the greater reason-generating force of harm. Further, it goes a fair distance in explaining the second asymmetry between harming to prevent greater harm and harming to provide a great benefit. Given the seriousness of harm, it becomes more evident why it might be permissible to engage in harm-minimizing behavior but not to impose such a condition in order to bring about a mere benefit.

It also offers resources to explain the third asymmetry between the moral significance of first-person election and third-person impositions of harm. Since this account identifies harm as a condition characterized by clash or severe estrangement from the will, then consent to harmful conditions would understandably change their significance. Consent involves active engagement with the features that generally make the condition harmful. Electing to accept a harm to achieve a purpose is a way of incorporating the condition into the larger narrative of one's life, thereby reducing much of the distance and clash that characterizes harm. Consent engenders a connection to the condition, but some aspects of the condition may still retain some harmful qualities; that the pain serves a chosen purpose reduces its significance but may not eliminate entirely the sheer awfulness of the pain—it remains a form of harm, even though chosen.

For this move to satisfy some of the desiderata, it will have to be reasonable for someone to elect to undergo harm to achieve a purpose or receive a benefit. We need to explain why the priority of harm does not exert the same hold in the first-person case to render the choice irrational. One answer is that one cannot build a distinctive identity and lead a good life just through deliberate efforts to avoid harm. Building a distinctive life requires seeking and pursuing relationships and projects, and this often involves seeking and choosing pure benefits. Indeed, achieving some of the important goods in life, both experiential and relational, demands exposing oneself to risks and direct experiences of harm (albeit within the appropriate personalized context and at the moments one is prepared to learn from them or appreciate their consequences). Achieving intimacy and sharing one's autonomous powers with others, for example, requires forms of vulnerability that may

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

instead sour and transform into dependence rather than sharing and collaboration, or other forms of failure and significant disappointment and emotional loss. And, of course, we may have to experience some of these sorts of harm to learn how to engage in intimacy well and how to appreciate its achievement.

From the first-person point of view, it is reasonable to place some priority on this larger process of self-definition, and to regard the process, its targets, and its products as more than merely pure benefits. The active pursuit of projects constitutes part of one's effort at self-definition and the achievement of a life suited to oneself; as I argue above, although it cannot be one's primary reason for pursuing one's particular projects, when they go well, their selection, pursuit, and integration into one's life have the collateral side effect of partly removing one from general conditions of passivity and alienation that constitute harm. This partly renders the decision to elect harm in pursuit of a project or relationship comprehensible and its significance different from mere bestowal by a third party. So third parties might balk at imposing harm on another or place a high priority upon their alleviating another's harm consistent with an agent reasonably opting to undergo the harm herself in order to further a project.

Hence drawing a connection between autonomy and harm may help to justify respect for the exercise of autonomy that results in self-inflicted serious harm. The moral significance of harm derives in large part, I believe, from the fact that the condition that strongly clashes with the person's will has the status of foreigner or invader. With self-inflicted harm, this characterization is less apt or is at least substantially complicated. The fact that it has been willed or is a by-product of one's own agency may reduce the extent to which the condition represents a harm at all but may also reduce its significance. The impetus to remove it is lessened because by virtue of its being the product of an agent's will, it is an element of his authorship; its origins are not foreign or alien. Unlike many other harms, it cannot fully be understood as an invasion of the person because it can intelligibly be thought to belong to him.

Moreover, averting the greater harm by constraining the agent's express and deliberate will seems inconsistent with the justificatory rationales underlying both the priority on harm and the recognition of autonomy rights. Intruding on a resistant agent's domain to prevent harm to her conflicts with what drives the priority on harm. Impeding the agent's will and forcing the agent to live according to the direction of another's will may prevent harm of one sort to the agent but introduces a serious, substantial foreign intrusion into her life. Although the condition imposed is not one the agent necessarily rejects in itself (for it averts greater harm to her), it is achieved by wresting control from her and subjecting her to the lights of another agent's will. This introduces alienation that is comparably harmful or at least of comparable significance to harm.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

## D. Belief-Mediated Distress

Before concluding, let me turn to a category of hard and interesting cases. In an important, original, and incisive discussion in *The Realm of Rights*, Judith Jarvis Thomson draws attention to a group of mental states she labels as "belief-mediated distress."[49] These are states such as anxiety, fear, or offense—states of distress that arise in response to a belief held by the sufferer. I may suffer anxiety because I believe that you may have failed to turn off the oven, fear because I believe that the masked man is threatening me, or offense because the masked man voiced sentiments I find shudderingly repellent. Are such states harms? Unlike states such as nausea or pain, as Thomson points out, belief-mediated feelings are ones that may be rational or irrational. Further, one may have rather direct responsibility for whether one has these feelings and with what intensity.[50]

Thomson denies that belief-mediated distress has the moral significance of other states of distress that occupy a status in her account analogous to harm, such as pain or nausea.[51] For Thomson, if there is a harm (or a morally comparable form of distress), then there is a claim that it not be caused. Thomson plausibly contends that one does not have a *claim* as such that others not cause one offense by living according to convictions we reject; nor do we have claims against suffering highly irrational forms of belief-mediated distress, such as the profound grief a bizarrely fastidious person might feel when her carpet is soiled.[52] Because there is no claim, therefore, suffering offense or irrational belief-mediated distress could not be a harm (or its functional equivalent).[53]

The problem extends further, she claims, to rational belief-mediated distress more generally. The fact that a claim infringement might also cause belief-mediated distress does not render the claim any more stringent, whereas standard harms do render claims more stringent. For instance, Thomson contends that in choosing between two claim-infringing actions, such as whether to trespass on A's land or B's land, whether one will cause one party belief-mediated distress does not supply an additional consideration against one such option. In her example, it is not a reason to prefer trespassing against A that B would suffer emotional distress that A has steeled himself against.[54]

I agree with Thomson that we lack *claims*, in her sense, that others not cause us offense by living in ways of which we disapprove or by voicing political sentiments we find abhorrent. So, too, we lack claims, in her sense, that

49. THOMSON, REALM, *supra* note 6, at 250.
50. *Id.* at 253.
51. *See supra* note 34 and accompanying text. *See also* THOMSON, REALM, *supra* note 6, at 253–259.
52. THOMSON, REALM, *supra* note 6, at 254.
53. *Id.* One might have a claim that another not engage in threatening act that gives rise to the belief and the distress, but the claim does not emanate from the harm of distress but from something else that is wrong about the threatening itself.
54. *Id.* at 254–255.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

others not cause us some exaggerated forms of irrational belief-mediated distress in response to everyday behavior, such as the paralytic fear I have when you blink in my direction.[55] And surely we lack claims, in her sense, that others not cause us appropriate guilt or other uncomfortable but morally apt emotions by drawing attention to morally relevant facts and arguments. My agreement, however, is qualified. We may well have claims that others not cause us even irrational distress *merely for sport* or for other poor reasons, for example, to bully or humiliate us. Similarly, we may have moral claims that others not cause us offense by insulting us as a form of noxious subordination. (Of course, that we may have such claims does not show that these forms of distress are types of *harms*.)

Unlike Thomson, though, I do not regard the question of whether something is a harm is so closely tied to whether we have full-fledged claims that others not cause it.[56] For some conditions may be harms that generate reasons for action but those reasons may not necessarily rise to the level of claims. If we focus only on states that give rise to claims on others, we may overlook important moral phenomena of intermediate moral significance. Further, if we attend too intently to those harms that generate claims that others not cause them, we may neglect the category of self-inflicted harms. That category seems particularly salient with respect to belief-mediated distress. It seems *possible* that I may inflict a harm on myself (even if I do not *choose* to harm myself deliberately) without violating a claim. Classifying some sorts of belief-mediated distress as self-inflicted harm rather than not as harm at all may represent a different way to respond to our frequently bearing responsibility for these mental states. Indeed, I suspect that one reason that belief-mediated distress is such an odd and difficult category is that many forms of belief-mediated distress seem to constitute, in whole or in part, self-inflicted harms, loosely speaking.

In some cases, such as that of appropriate guilt, one's actions and one's correct moral judgments may be the underlying causes, even if the temporally proximate event that sparks the guilt is another's reproach. In other cases, such as when I feel terrified when I see that you wear red, your wardrobe decisions may cause my belief that you wear red, but my irrational reaction to red clothes causes my distress. That my proclivity to overreact and that the emotional states that result are deeply irrational and perhaps avoidable or curable means they are in an important sense self-generated, but they may still be harms. Further, the bearer of this sort of self-inflicted harm, one involving epistemic and normative misjudgments, may have special first-personal reasons to address that harm that others lack. So I am more open to the idea that some forms of belief-mediated distress may constitute harms, although they may more frequently be self-inflicted and not provide the basis for full-fledged claims.

55. *Id.* at 254.
56. *See supra* note 6 and accompanying text.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

The main source of the disagreement between Thomson's skeptical conclusion that belief-mediated distress cannot be a harm and my greater openness toward the view that such distress may constitute a harm traces to the divide between her claim-centered approach to what constitutes harm and my broader, reason-centered approach. Still, the arguments she offers lay bare many complexities about belief-mediated distress that require further investigation on any approach. In what follows, I say a little more to motivate my sense that some forms of belief-mediated distress do constitute harms, even though we bear some responsibility for whether we are in such states. I also discuss, however, why Thomson's arguments suggest that offense in particular may fall outside the category of harm.

I find it difficult to resist the idea that reasonable fear and grief, say in response to a loved one's death, are harms. The sort of overwhelming invasion of grief or fear into one's consciousness, as well as its regular occasioning of mental (and sometimes physical) paralysis, seems as morally serious as pain or nausea. Grief and fear can disrupt agency in ways similar to non-belief-mediated states such as pain or nausea. Perhaps for similar reasons, some cases of persistent, reasonable anxiety are harms. That these states are mediated by our beliefs, including our values, does not seem to make *all* of the difference. We have moral reasons to try to alleviate the suffering of those experiencing grief and fear. Further, although it is largely overdetermined, we also have moral reasons to avoid causing the facts underlying the reasonable distress,[57] for example, we have reasons to avoid causing deaths and frightening people in part because of the resultant grief and fear.[58] For instance, that an autonomous adult's jaywalking habit causes his parent reasonable anxiety or fear constitutes an additional reason to refrain from jaywalking, above and beyond the reasons provided by the perhaps quite slim danger of his activity. Further, the grief one's friends and family would suffer at one's death gives one an additional reason to resist the temptation to jaywalk.

Even some forms of irrational fear can create reasons. Suppose that to attract attention to my business, I craft an awning over the storefront that sits at a rather rakish but not especially artistic angle. The awning is in fact structurally sound, and a pitched political battle has made that fact well known. Still, despite their knowledge of the security of the installation, a few cannot shake the feeling that it is about to crash to the sidewalk, and they experience waves of irrepressible fear when they walk by. Their persistent fear seems like a harm; whether it gives them a claim against me that I alter the awning, I am unsure, but the fact that it provokes fear gives me

57. Greg Keating, in a thoughtful discussion, argues for considering some forms of emotional distress as forms of harm, at least for legal purposes. Gregory C. Keating, *Is Negligent Infliction of Emotional Distress a Freestanding Tort*, 44 WAKE FOREST L. REV. 1131 (2009).

58. I would not suggest, however, that the harm of grief provides standing reasons not to inform someone of a death. How we avoid and alleviate these harms must be appropriately responsive to the values that underlie them.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

at least some moral reason to advertise my business through means other than radical architecture.

Thomson's seemingly contrary case of the irrelevance of one party's distress to the choice between trespassing on the land of the delicate flower versus trespassing on the land of the trained stoic may have a limited scope. Although a persuasive case, it may not provide a clear counterexample to the thought that much belief-mediated distress may generate moral reasons of the sort generated by harm. Her case might be explained by the injustice of allowing a person's successful efforts to reduce the degree of his emotional vulnerability to a wrong to constitute a reason that very person should suffer a wrong instead of another who did not undertake such measures.[59] Consider another example involving reasonable belief-mediated distress without that feature. Suppose that to raise money for medication for a dangerously ill family member, one had to choose between stealing an inherited brooch from a home versus stealing a brooch of identical financial value from a jeweler, with equally significant financial consequences for the victims. The former theft would also cause the owner emotional distress because of the brooch's sentimental value. Its loss would reignite grief about the passing of the brooch's original owner; whereas the theft of the latter would not cause distress because this jeweler has no special emotional attachment to her wares. Here, the distress does seem a reason to prefer to infringe the claim of the jeweler instead of that of the heir. The heir may lack a further *claim* that he not be caused this distress on top of the claim he has against the theft; still, the theft would harm him in a morally relevant way that it would not harm the jeweler.

Despite my sense that some forms of belief-mediated distress constitute harms (in my sense), I endorse Thomson's skepticism about offense constituting harm or giving rise to the reasons associated with harm.[60] How, then, does offense differ from grief or fear, which seem to me to be eligible candidates for harm? The difference cannot be that grief at the loss of a loved one is rational but offense is irrational. Some offense is rational and justified, whereas some sorts of irrational fears seem to be harms. Nor is offense necessarily self-inflicted. Think of the offense an audience experiences in response to a sexist or racist speaker. Nor can be said exactly that the difference is that one thoroughly identifies with being offended in a way that is untrue of grief. As with grief, most would prefer that the action or state giving rise to offense not happen, yet they endorse the judgments and

59. Frances Kamm makes a related point that some conditions that normally give rise to reasons (and claims) may be irrelevant to certain choices and therefore fail to render a claim more stringent. For instance, if we must choose whether to divert a trolley onto one person rather than another, where the diversion will cause either victim to lose a leg, it should seem irrelevant that one victim would also suffer a sore throat (even if usually claims restrict us from causing sore throats). Her analysis suggests that some states may give rise to reasons in some circumstances but may be overwhelmed or rendered irrelevant in others. *See* FRANCES MYRNA KAMM, 1 MORALITY, MORTALITY: DEATH AND WHOM TO SAVE FROM IT (1993), at 144–164.

60. THOMSON, REALM, *supra* note 6, at 257–259.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

values that place them in the relevant emotional state. They would rather not be in the emotional state because they regard it as terrible that its object is true, but given that its object is true, they want to be the sort of person who experiences these emotional reactions.

So what could the difference be if we acknowledge that fear, grief, anxiety, and offense may all be rationally grounded states of belief-mediated distress? I do not have a good account. Perhaps rational offense and irrational offense (or reasonable and unreasonable offense) are on all fours with rational fear and irrational fear, respectively. But I am tempted to think that offense may differ for the following reason. The sort of offense a liberal theorist is reluctant to recognize as a form of harm, for example, offense at one's radical ideas or sexual orientation, arises in response to the specific content of actions and judgments for which it is an essential component of autonomy rights that individuals enjoy the freedom to make their own decisions about them. Further, that freedom is fully meaningful when such actions and judgments are decided largely "on their own merits" and not because of direct or indirect social pressure to think or do otherwise.[61] To regard even rational and justified offense as harmful and to think it might give rise to strong reasons to avoid it or alleviate it qua harm would generate the very sort of social pressure to avoid the underlying thoughts or ideas that is in tension with our recognition of autonomy rights in this domain. (Of course, there may be other reasons not to cause offense, for example, to begin a cordial relationship or avoid conflict.) Whereas, the states of the world that give rise to (rational) fear and grief are not ones that arise in this way. Acknowledging reasons to alleviate fear and grief and to avoid causing the objects of this distress do not, as a standard matter, stand in direct tension with the practice or foundation of central autonomy rights.[62]

If we were to take this line, anxiety provoked by others' actions would appropriately represent a hard case, falling somewhere in between offense, on the one hand, and fear and grief on the other. In some cases, the behaviors to which anxiety represents a response resemble those that may give rise to offense but arise in the context of special relationships. Those special relationships may permissibly allow one to exert some pressures in domains where strangers, acquaintances, and colleagues may not. Further, the belief giving rise to the distress need not involve disagreement about the value of another's self-regarding activity. It may relate to a concern that some value reasonably considered as shared is in jeopardy, as when one business partner experiences nagging but appropriate anxiety about whether his counterpart has mailed their joint tax return, given the latter's

61. *See* my related discussion in Seana Valentine Shiffrin, *Egalitarianism, Choice-Sensitivity, and Accommodation*, *in* REASON AND VALUE: THEMES FROM THE WORK OF JOSEPH RAZ 270–302 (Philip Pettit et al. eds., 2004).

62. A somewhat similar story might be told about appropriate guilt. To recognize it as a harm associated with moral reasons not to evoke it would be in tension with the grounds that render it appropriate.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

chronic tardiness; here, the anxiety seems like a harm that provides the business partner an additional though not the primary moral reason to turn punctual. For a close case, we might return to the parent who is anxious about his adult child jaywalking; that anxiety may be a harm, whereas the nosy neighbor's anxiety about the jaywalking may seem more like offense than (even self-inflicted) harm.

Thus my tentative conclusion is that some forms of belief-mediated distress, such as fear or grief, may indeed be harms. In many cases, they may not give rise to full-fledged claims against others not to cause them but they regularly give rise to reasons in others to avoid causing their underlying objects and to alleviate them when present. In some cases, particularly those of substantial irrationality or appropriate guilt and shame, we may think that these harms are self-inflicted and that they primarily generate reasons *for the sufferer* to avoid, prevent, or alleviate them. That fact that those harms are self-inflicted may affect the strength of the reasons others have to avoid or alleviate them as well as the sort of avoidance or relief others should offer. Rather than avoiding the actions that cause the beliefs giving rise to the distress, the more appropriate form of moral aid may be to offer reasons, counsel, opportunities for remorse and repair, and, with respect to irrational distress, perhaps other forms of therapy to address the underlying irrationality.

Finally, and much more tentatively, not all forms of belief-mediated distress may be harms. Offense and some sorts of anxiety may not be harms because they could not reliably give rise to the sorts of reasons to avoid or alleviate them consistent with the conditions of meaningful autonomy; those conditions, I am suggesting, bear a close relationship to harm and its moral significance.

## V. CONCLUSION AND FURTHER CHALLENGES

I have argued that significant problems with comparative interest-setback accounts should encourage us to rethink our understanding of harm. Although common, appeals to the doing/allowing distinction are ill suited to vindicate asymmetries between harms and benefits and put the wrong sort of pressure on that distinction. Given the prominence of consent and the will in these asymmetries, it seems likely that some connection to the will should surface in our theories of harm. An account that isolates the passivity that harm involves and highlights the importance of one's experience and one's life being rendered one's own has promise, especially because it holds out the possibility of a more integrated account of autonomy and our sense of the special negative significance of harm.

Still, such an account faces substantial challenges that I do not resolve here. It would require a more detailed and precise account of what counts as passivity, what counts as an impediment to the will, and what counts as an

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press

objectionable chasm between the will and one's experience. Further, such an account would need to make further progress on issues about how harm and other foundational moral standards relate.[63] It would need to confront moral harm in depth, for example, on the one hand, whether and in what respect living an evil life may constitute, at least in part, a self-inflicted harm to its possessor, because an evil life represents a sort of alienation from objective moral standards, and on the other hand, whether the destruction or frustration of evil projects should count as harms or not. It would also require a substantially more elaborated account of whether, when, and why disappointments, frustrated desires, and minor obstructions of one's projects count as harm and when they should be thought of as mere setbacks of interest; that is, a more developed account would be necessary to ground the difference between losses and failures that create a chasm between a person's will and her experience versus mere preference frustration, which one might often resist but which we would not count as harm (at least not if we also retain a sense of harm's priority).[64] Tied up with these issues are the questions of *how* one's consent makes a difference and how much of a difference it can make.

Although these challenges are large, they seem like the correct ones to face. The obstacles they pose do not seem so great that they would justify returning to the standard models, models that cannot even fully make out these rather familiar moral questions as sensible.

63. This issue of the degree to which what "harm" encompasses incorporates or may conflict with other foundational normative standards is fleetingly broached in the discussion above of comparative accounts' implication that equality norms (of justice) are in necessary tension with harm's priority, at note 33 and accompanying text, in the idea that conditions caused by rights violations may be harms even if the condition as such would not otherwise be, and also in the discussion of belief-mediated distress in Section IV.D.

64. The answer may depend on the source and strength of the obstruction and how significant the endeavor is to the agent. Although certain significant discordances between the will and one's experience, life, or circumstances are the key components of harm, complete achievement of what one wants is not necessary to avert harm. Many preferences may go unfulfilled without making one's life foreign and without, in contrast with pain, presenting an intrusive, noxious, or importantly alien element into one's direct consciousness. Of course, persistently present preferences that one disapproves of or otherwise disidentifies with but that resist efforts at removal may themselves represent harm. It is their continued presence, rather than their nonfulfillment, that is harmful. On the other hand, persistent and widespread frustration of all or very many of one's preferences may deposit one within circumstances that render one's life significantly foreign and to be endured, watched from within. Likewise, not all obstacles to one's projects will represent harm, or at least, not overall harm. The ability to choose one's endeavors and to participate in efforts to promote them matters here, as does the fact that many endeavors reasonably involve risk and competition. To will them is to will participation in an endeavor whose success may depend upon the achievement of excellence through collective striving and competition. One may not desire frustration, but it will occur as part of a process to which one has committed and whose fruition more integrally depends upon competition for the achievement of its aims. With some projects, realistic appreciation of their nature means that one's reasonable pursuit of them is not necessarily aimed at success or achievement of their aims but at making a good-faith effort, so that some forms of failure are, at bottom, consistent with what one has willed. *See also supra* note 43.

https://doi.org/10.1017/S1352325212000080 Published online by Cambridge University Press