Daniel Immerman

Harm, baselines, and the worse than nothing account

Harm is one of the central concepts of ethics so it would be good to offer an account of it. Many accounts appeal to a baseline: they say that you harm someone if you leave them worse off than in the baseline case. In this paper, I draw some lessons regarding what counts as an appropriate baseline and explore what these general lessons reveal about the nature of harm. In the process of so doing, I argue that a certain rarely-discussed account of harm – the worse than nothing account of harm – does a particularly good job at identifying a baseline. This account says you harm someone if you leave them worse off than if you had done nothing to them.

Harm is a key concept in ethics, so it would be good to provide an account of it. Many accounts of harm share a common feature: they use a baseline to determine whether something counts as a harm.[1] One way to evaluate these accounts is to explore the baseline they provide. Does it allow them to correctly classify cases? Is it fruitful in resolving puzzles? Is it principled? Does it help us better understand the nature of harm?

This paper looks at a number of these accounts and focuses in on their baselines. In critically examining them, it draws some general lessons regarding what sort of baseline we should be looking for. In addition, it explores what these lessons reveal about the nature of harm.

Not all accounts fare equally well. I will argue that a certain rarely-discussed account of harm – the worse than nothing account of harm – does a particularly good job at offering a baseline.[2]

---

[1]This set of accounts is commonly acknowledged to be the most promising accounts of harm. See e.g. [Bradley, 2012, 396], [Feit, 2015, 361], [Norcross, 2005, 150]. For criticism of alternative accounts of harm, see e.g. [Boonin, 2014, 56-102], [Bradley, 2012, 398-405], [Carlson et al., Forthcoming], [Johansson and Risberg, 2019, 351-352], [Klocksiem, 2012, 287-8], [Rabenberg, 2015, 2-8], [Woollard, 2012].

[2]Roughly speaking, the worse than nothing account says that your action harms someone if it leaves them worse off than if you hadn't done anything to them. While there are some older accounts in the literature that bear a resemblance to the worse than noth-

At some points in this paper, I will be offering counterexamples to certain accounts of harm. Some of these echo ones that have been proposed by others. So the contribution of this paper is not in providing these examples. Rather, it is in seeing what they reveal regarding what baseline we should be looking for and what this in turn reveals about the nature of harm.

In addition, as I just noted, I think these lessons speak strongly in favor of a particular account of harm – the worse than nothing account. So another contribution of this paper is to show that this account accommodates these counterexamples and the lessons they provide. Indeed, I shall go a step further, arguing that not only can this sort of account accommodate these lessons, but that it is the only sort of account that can.

This paper has five sections. The first clarifies terms. The next three discuss rivals to the worse than nothing account and raise problems concerning their baselines. The final section introduces the worse than nothing account and argues that it can overcome the problems that face its rivals.

1 Clarifying terms

Let me start by clarifying terms. First: harm.

As with most in the literature, my main focus is on ultima facie harm, as opposed to prima facie harm.[3]

Some offer accounts of harmful actions, while others offer accounts of harmful events.[4] The second category is broader. For example, suppose a tree branch falls on me, injuring my shoulder. This is not an action; tree branches are not agents. Nonetheless, it harms me. Given that there are

---

ing account (see e.g. [Bennett, 1995, 113-119], [Donagan, 1977], [Howard-Snyder, Summer 2002], [Kagan, 1989, 97-8]), it is relatively neglected in recent years, with one exception: a recent paper by Daniel Immerman [Immerman, 2021]. His paper focuses on the account's ability to solve a problem called "the preemption problem" whereas this paper focuses on its ability to identify a baseline.

[3]See e.g. [Bradley, 2012, 393-4], [Feit, 2015, 361], [Feit, 2017, 810], [Hanna, 2016, 251], [Immerman, 2021, 4], [Johansson and Risberg, 2019, 352], [Norcross, 2005, 150], [Purves, 2019, 2630]. By contrast, some focus on pro tanto harm, e.g. [Purves, 2016, 68].

[4]For those who take the counterfactual comparative account to be an account of *actions*, see e.g. [Boonin, 2014, 52], [Norcross, 2005, 150] for those focused on events, see, e.g. [Bradley, 2012, 391], [Feit, 2015, 361], [Feit, 2017, 810], [Hanna, 2016, 251], [Johansson and Risberg, 2019, 351], [Klocksiem, 2012, 285], [Northcott, 2015, 151], [Purves, 2016, 68], [Purves, 2019, 2631], [Rabenberg, 2015, 8-9].

harmful events that are not actions, it seems best if an account of harm can accommodate them.

I will be focused in this paper on "comparative accounts of harm". These account for harm by means of a contrast case. In particular, they say that an action (or event) harms someone if it leaves them worse off than in this contrast case. I refer to this contrast case as a "baseline".

Several notes about these accounts.

First, strictly speaking, these accounts are accounts of harming. That is, they determine when a particular action or event qualifies as a *harming* of someone.[5]

It will sometimes be useful to have a term for this action or event: the one of which we are asking "is it a harming?". I will call it "the main event". So, for example, if I'm trying to decide if my cat's scratching me counts as harming me, the cat's scratching me is "the main event".

Second, these accounts are relational: they concern whether an action or event qualifies as a harming of *someone*. After all, it is possible for the very same action to harm one person and help another. For instance, if a candidate makes a particularly devastating point about their opponent in the midst of a political debate, they simultaneously help themselves and harm the other candidate.

It will be useful to have a term for this someone, the one of whom we are asking: "were they harmed?". I will call them "the recipient". So, for example, if I'm trying to decide if my cat's scratching me counts as harming me, then I am "the recipient".

2   The temporal account

As I noted in my introduction, I will draw general lessons about baselines and harm by working through various comparative accounts. Let me start with the following:

> *The Simple Temporal Account.* A's doing x harms B just in case it leaves B worse off than before.

---

[5]Compare: some also wish to give accounts of "harm" understood as a state of affairs. So for example, those endorsing the causal theory of harm say that you harm me (in the action or event sense of "harm") just in case you cause me a harm (in the state of affairs sense of "harm").

To illustrate the account, let us use an example. Suppose my cat scratches me. Did she harm me? According to the simple temporal account, we should look at my state prior to her scratching me. I was doing better prior to her scratching me than after she did so, so the simple temporal account correctly classifies this as a harm.

I am about to criticize this account, but before I do so, I wish to note that it has some seeming attractions. Perhaps the biggest one is that, unlike the accounts I will be considering in later sections of this paper, it does not invoke counterfactuals. Why is this an advantage? Counterfactual accounts in other domains are rife with problems. It wouldn't be surprising if some of these problems carry over to counterfactual accounts of harm. An account that avoids counterfactuals will sidestep these potential problems.

Unfortunately, the baseline provided by the simple temporal account runs into problems. To illustrate, here's an example:

> SLOW DECLINE. I have a disease that is making me feel worse and worse. My doctor gives me a treatment. While I continue to decline, I would have declined faster without the treatment.[6]

Intuitively, the doctor's treatment benefits me. And this is so even though I am doing worse after the treatment than I was beforehand.

The simple temporal account fails to handle this case. It says that giving me the treatment harms me. After all, I am doing worse after I get the treatment than it was before I began it. But this is the incorrect verdict; it was the disease that harmed me, not the treatment.

Let me develop the point further. The reason my prior state is not an appropriate baseline is that there are multiple factors at play in this scenario: the disease and the treatment. When I ask if the treatment harmed me, I want to focus in on the evaluating the treatment as opposed to the disease. But the way the simple temporal account sets up its baseline doesn't allow us to treat these separately. Instead, our baseline – my state before I started the treatment – differs in multiple ways from the state we are comparing it to – my state after the treatment is complete.

We can see how to provide a better baseline by noting a related point regarding medical practice. If researchers are trying to decide whether a treatment is helpful or harmful, they do not perform a simple before and after comparison. Instead, they compare two sets of patients: patients that

---

[6]For some other related criticisms, see e.g. [Boonin, 2014, 59].

got the treatment and those that didn't. This way, they can isolate the effects of the treatment while holding other factors constant.

In sum, here is our first lesson:

Lesson 1: A good baseline should hold external factors constant.

The simple temporal account fails because it does not hold external factors – the illness – constant.

What factors count as external? We can answer this question via appeal to causation, counterfactuals, or a combination of the two.

To illustrate, return to the case of SLOW DECLINE. The treatment did not cause my illness. And whether or not I received the treatment, the illness would have persisted. So we can appeal to causation or to counterfactuals (or to both) in explaining why the illness was external to the treatment.

Of course, this is not to say that illnesses are always external to treatments. For example, suppose that I am sick, my doctor gives me a treatment, and this in turn causes a secondary illness.

Here, we do not get the verdict that the secondary illness is external to the treatment. First, it is caused by the treatment. Second, if I hadn't gotten the treatment, I wouldn't have gotten the secondary illness.

Because the secondary illness is not external to the treatment, we should not hold it constant when generating a baseline. In other words, suppose we are examining this second case, trying to determine if the doctor's treatment helped or harmed. If so, when we generate a contrast case, it should be one where I got neither the treatment nor the secondary illness.

I emphasized early on in this paper that harm is an important concept in ethics. So it's worth noting that this exclusion of external factors makes sense from the perspective of morality. If I wonder whether an action of yours was right or wrong, it makes sense to exclude such factors. After all, they are not part of your action in any meaningful way. So, for example, suppose we wished to judge the doctor's actions in SLOW DECLINE. If so, it would be unfair to blame the doctor for the illness. The illness is not part of the doctor's action, nor is it an effect of it. The doctor's contribution is positive; the doctor helped the patient, even though the patient continued to deteriorate, thanks to the illness.

As I noted, one can try to characterize the notion of an external factor using causation, counterfactuals, or a combination of the two. Let me briefly

explore the first option, in which external factors are characterized entirely through causal notions.[7]

One attraction of doing so is that would allow us provide a modified version of the simple temporal account that avoids our earlier counterexample:

> *The Causal Temporal Account.* A's doing x harms B just in case it causes B to be worse off than B was before.[8]

This new account avoids the problems with SLOW DECLINE. After all, the treatment does not cause the patient to deteriorate.

In SLOW DECLINE, causation and counterfactuals run together. More interesting is a case where they come apart:

> NEAR ACCIDENT. You are crossing the street and a distracted driver is about to hit you. I push you out of the way, causing you to fall and skin your knee.

In this case, I cause you to be in a worse state of affairs than you were previously. Before I pushed you, you were happy and carefree. After I pushed you, you were in pain and your knee was bleeding.

Counterfactually speaking, by contrast, you are in a better state. Yes, pushing you startled you and caused some injury to your knee. But it left you better off than you would have been had the driver crashed into you.

Did I harm you? It seems to me that the answer is "no", at least if we are considering an all-things-considered sense of "harm". All things considered, I benefitted you – I saved your life.[9]

In short, when we look at cases that pull apart causation and counterfactuals, we see that we cannot account for external factors merely in terms of causation.

Again, this makes sense from the perspective of morality. When it comes to deciding how to act, we cannot simply ignore how others will behave. There is a morally significant difference between pushing someone out of the way of a car and pushing them for no good reason, even if the causal effects of one's actions are identical. The idea that in the first case, one is helping and

---

[7]Thanks to several anonymous referees for encouraging me to do so.

[8]See e.g. [Boonin, 2014, 57].

[9]For similar cases and verdicts, see e.g. [Carlson et al., Forthcoming, 13], [Immerman, 2021, 18]. The Carlson Et. Al. paper also has a number of other criticisms of the causal temporal account.

in the second case, one is harming seems well placed to explain this moral difference.

3   Worse than the alternative account

In the last section I argued that we should not endorse the simple temporal account and that if we wish to hold external factors constant, then our account will most likely need to invoke a counterfactual. If we wish our contrast case to involve a counterfactual, we will commit to what is called a "counterfactual comparative account". As its name suggests, this is a special sort of comparative account: one which uses, as its baseline, a counterfactual.

If we take this step, we will be in good company. The counterfactual comparative account is currently the most popular account of harm.[10] It is intuitive,[11] has a good deal of explanatory power,[12] and is able to dodge a number of objections that other accounts face.[13]

There are multiple versions of this type of account, depending upon which counterfactual we select. Here is the most popular version of the counterfactual comparative account:

> *The Worse Than The Alternative Account Of Harm.* A's doing x harms B just in case it leaves B worse of than if A hadn't done x.[14]

Let me give an example to help illustrate. We'll return to the case of my cat scratching me. Let's suppose that if she hadn't scratched me, she would have watched the birds through the window. Assuming so, the worse than

---

[10]See e.g. [Bradley, 2012, 396], [Feit, 2015, 361], [Norcross, 2005, 150].

[11]For those who note its intuitiveness, see e.g. [Boonin, 2014, 52], [Feit, 2015, 362], [Feit, 2017, 810], [Johansson and Risberg, 2019, 351], [Klocksiem, 2012, 285].

[12]For those who emphasize its explanatory power, see e.g. [Boonin, 2014, 52], [Bradley, 2012, 397], [Feit, 2015, 362], [Hanser, 2008, 434-7], [Johansson and Risberg, 2019, 351], [Klocksiem, 2012, 285], [Purves, 2016, 69], [Purves, 2019, 2631].

[13]For criticism of alternative accounts of harm, see e.g. [Boonin, 2014, 56-102], [Bradley, 2012, 398-405], [Johansson and Risberg, 2019, 351-352], [Klocksiem, 2012, 287-8], [Rabenberg, 2015, 2-8], [Woollard, 2012].

[14]For formulations of this sort of account, see e.g. [Boonin, 2014, 52], [Bradley, 2012, 396], [Feit, 2015, 361], [Feit, 2017, 810], [Feldman, 1991, 215-6], [Hanna, 2016, 251], [Hanser, 2008, 423-4], [Johansson and Risberg, 2019, 351], [Klocksiem, 2012, 285], [Norcross, 2005, 150], [Purves, 2016, 69], [Rabenberg, 2015, 8].

the alternative account says that her scratching me counts as a harm just in case it leaves me worse off than if she had watched the birds through the window.

As I just noted, this account is the most popular version of the what is itself the most popular account of harm: the counterfactual comparative account. Perhaps unsurprisingly, it has a number of things going for it. Most relevant to our discussion thus far is that it has no problem accommodating SLOW DECLINE. It says that to figure out if my doctor's treatment harmed me, we should compare her treating me with what would have happened if she hadn't treated me. If she hadn't treated me, I would ended up worse off, so the account delivers the correct result that her action is not a harm. Similar things can be said for NEAR ACCIDENT.

That said, and perhaps unsurprisingly, the worse than the alternative account does face challenges. One of the biggest ones is called "the preemption problem".[15] It centers around cases in which something bad happens, and if the bad thing hadn't happened, something even worse would have happened.

To see this, let us look at an example:

> KNIGHT. Bobby Knight gets mad at a philosopher, Phil, in response to a perceived slight. So he chokes Phil. Luckily for Phil, Knight is in anger management. It's taught him to better control his behavior. He choked Phil because he applied anger management techniques. If he hadn't, he would have dismembered Phil. [Hanna, 2016, 263].[16]

The worse than the alternative account says that an action harms someone if it leaves them worse off than if it hadn't occurred. So it instructs us to compare the choking with what would have happened had the choking not

---

[15]For further discussion, see e.g. [Boonin, 2014, 62], [Bradley, 2012, 407-8], [Feit, 2015, 363], [Hanser, 2008, 433-4], [Immerman, 2021], [Johansson and Risberg, 2019], [Klocksiem, 2012, 287], [Norcross, 2005, 165-6], [Northcott, 2015, 159], [Rabenberg, 2015, 10], [Roberts, 2007, 271], [Thomson, 2011, 446-7]. For places where people say this is one of the most serious problems facing the counterfactual comparative account, see e.g. [Bradley, 2012, 398], [Feit, 2015, 363], [Johansson and Risberg, 2019, 353].

[16]The case is originally due to Alastair Norcross [Norcross, 2005, 165-6]; I'm quoting a condensed version of it from Nathan Hanna. For discussion of it and similar cases, see e.g. [Boonin, 2014, 62], [Bradley, 2012, 407-8], [Feit, 2015, 363], [Hanser, 2008, 433-4], [Johansson and Risberg, 2019], [Klocksiem, 2012, 287], [Norcross, 2005, 165-6], [Northcott, 2015, 159], [Rabenberg, 2015, 10], [Roberts, 2007, 271], [Thomson, 2011, 446-7].

occurred. The choking was bad, but the dismemberment is worse. This means that the account delivers the incorrect verdict that the choking was not a harm.

Again, let us try to dig deeper and see what has gone wrong with the baseline provided by this account. Here, the problem is as follows: our baseline replaced the main event (i.e Knight's choking) with something that was even worse (i.e. the dismemberment). The dismemberment does not seem like a "neutral" comparison point. Instead, its negative valence means that something could be better than it and still be very bad. Indeed, this is exactly what we find in the case of the choking.

This provides our second lesson:

> Lesson 2: A good baseline should replace the main event with one that is neutral.

What does it mean for an event to be neutral? Again, there is room for reasonable disagreement on the details (more on this in the next two sections), but here is a minimal version that everyone should be willing to get behind: when I act neutrally towards you, I neither help you nor harm you. After all, if our baseline involved me harming you, then I could leave you better off than it and still harm you, so long as I performed a more minor harm. Likewise, if our baseline involved me helping you, then I could leave you worse off without harming you; in particular, I could help you less.

It is clear, given the way that the worse than the alternative account is set up, that its baseline is not neutral. So, for example, when Knight dismembers Phil, this is not a neutral event, because it is a way of harming Phil.

As with the previous lesson, this lesson makes sense from the perspective of ethics. In trying to decide if my action was acceptable or not, our comparison point should lack a valence. If our comparison point involves significant moral wrong, then I can do somewhat better while still acting inappropriately. If our comparison point involves significant moral benefit, then I can do somewhat worse while still doing well by you.

It is perhaps worth noting that if we supplement Lesson 2 with my minimal characterization of neutrality, it will be rendered circular. In particular, it will say that a good baseline should replace the main event with one in which I neither help you nor harm you. This is circular insofar as it references

harm in stating a condition of adequacy on an account of harm.[17]

I think this is probably fine. I am happy to grant that circularity can be problematic when it shows up in an account of harm. But here my goal is different; it is to provide adequacy conditions on an account of harm. Circularity in adequacy conditions on an account seems much less bad.

Compare: one might think that a condition of adequacy on an account of knowledge is that it preserves closure. That is, one might think that an account of knowledge should ensure that if S knows p and p entails q then S knows q. One could then criticize certain accounts of knowledge (e.g. sensitivity accounts) that violated closure.

The fact that this condition of adequacy references knowledge seems entirely unproblematic. Even though it references knowledge, it is still useful; it helps rule out certain accounts.

That said, I do think there are attractive ways of spelling out the notion of "neutrality" that do not reference harm. I will be discussing some of these in the next two sections.

4 Moral minimum comparison accounts

In the last section I said that there is room for disagreement on exactly what counts as neutral. Let us look into this some more. I will start with a key decision point: are we talking about a moral sense of neutrality or a descriptive sense?

Let me say more about each of these options, starting with the first. This first option is to spell out the notion of neutrality in moral terms. There are various ways in which we might do this, but for illustration let me focus on one, which I'll call "doing the moral minimum". If one does the moral minimum, one fulfills one's moral obligations but does not go above and beyond the call of moral duty.

To illustrate the moral minimum, it would be helpful to have some examples. First, negative examples. Here are some examples of actions that do not count as doing the moral minimum: choking Phil, viciously attacking Phil, shooting Phil. These do not count as doing the moral minimum because they violate an obligation.

Here are some further actions that do not count as doing the moral minimum: giving Phil all of one's money, donating a kidney to Phil, buying Phil

---

[17]Thanks to [name redacted] for raising this issue.

an all-expenses-paid cruise. They do not count as doing the moral minimum because they go above and beyond the call of duty.

Finally some positive examples: here are some examples of things that would have counted as doing the moral minimum: nodding calmly at Phil, ignoring Phil, wishing Phil a pleasant day.

As we have just seen, the moral minimum account of neutrality understands neutrality in a morally-laden way. As I noted, there is an alternative way of fleshing out the notion of neutrality. Instead of using moral language, we can flesh it out descriptively.

Again, there are a number of versions of this, but for the sake of example, here is one: doing nothing to the other person; just leaving them be.

Which of these is better when it comes to a baseline for accounts of harm? In this section I discuss approaches that spell out neutrality using moral language. Again, for the sake of concreteness, I will focus on an account that understands neutrality in terms of doing the moral minimum:

> *The Moral Minimum Comparative Account.* A's doing x harms B just in case it leaves B worse off than if A had done the moral minimum.[18]

It will be helpful to give an example to illustrate the account. We can use the example from last section. The moral minimum account deems Bobby Knight's action a harm. This is because his action leaves Phil worse off than if Bobby had done the moral minimum. If Phil had done the moral minimum, e.g. by casually nodding at Phil, Phil would have ended up better off.

Again, there are some reasons to think this account is worth taking seriously.

First, it does a nice job handling cases that prove tricky for rival accounts. So, for example, as we have just seen, it handles a key instance of the preemption problem. The preemption problem is commonly viewed as an extremely serious problem for counterfactual comparative accounts. Indeed, it is such a serious problem that it leads some to suggest that we should abandon counterfactual comparative accounts;[19] and perhaps even excise harm from our moral theorizing.[20]

---

[18]For similar accounts, see e.g. [Boonin, 2014, 61].

[19]See e.g. [Bradley, 2012, 398], [Feit, 2015, 363], [Johansson and Risberg, 2019, 364].

[20]See [Bradley, 2012, 410-11].

Second, this account seems to offer attractive verdicts for cases concerning neglect. For example, suppose that I neglect my pet dog. To many, this seems to be a case of harm. But this is a verdict that many counterfactual comparative accounts struggle with.

To illustrate, take the worse than the alternative account. This account says that to see if my neglect is a harm, we should compare it what I would have done otherwise. Suppose I'm a moral monster. Suppose that if I hadn't neglected my dog, I would have treated him even worse; suppose I would have abused him. This forces the worse than the alternative account to say that the neglect is not a harm; indeed, it seems to classify this as a benefit!

The moral minimum account, by contrast, seems to get this case right. In particular, it compares the neglect with the case in which I fulfill all my obligations. If I had fulfilled all my obligations, including my obligations to my dog, he would have ended up better off, so the moral minimum account says that in neglecting my dog, I harm him.

While this account has various attractions, it does face problems. First, it is limited in scope. In particular, it only applies to actions. After all, it's not clear what it would mean for tree branches to perform the moral minimum.

Is this a serious problem? It's not clear to me. When it comes to ethics, we care mostly about action. So if an account can handle all cases involving action, this seems perhaps to be good enough.[21]

Second, the moral minimum account will not be attractive to those who wish to account for harm without referencing morality. Perhaps, for example, you wish to account for moral obligation in terms of prima facie duties. Perhaps one of these duties is to do no harm.[22] If you then accounts for harm in terms of morality, your account will be circular.

Is this a serious problem? Again, it strikes me as not necessarily fatal; there are multiple reasons you might want an account of harm, and not all of these require that one account for harm in a way that doesn't reference morality.[23]

A more serious problem for the account centers around its baseline. Here is a case to illustrate the problem:

DOG ATTACK. I am out on a walk with my toddler, when an

---

[21]Of course, as I noted earlier, it would be even better if we could accommodate all harmful events and not merely those that are actions.

[22]See e.g. [Ross, 1930].

[23]For some examples, see [Bradley, 2012, 390].

angry dog comes running out of a nearby house. It is clear that the dog is about to attack my toddler. The only way to stop the dog from attacking is to throw a rock at it. I do so and hit the dog in the shoulder, injuring it. The dog yelps and runs away.[24]

Unfortunately, the moral minimum account gets this case wrong. It says that my action does not harm the attacking dog. After all, I do the moral minimum here; given my duties as a parent, we may suppose, I was obligated to act as I did and throw the rock. This seems like the wrong verdict; clearly, in hitting the dog with the rock, I harmed it.

Let us dig deeper and see what has gone wrong with the moral minimum account. The problem seems to be that in some cases, we can be obligated to harm. So for instance, in DOG ATTACK, given my duties as a parent, I was obligated to harm the dog so as to protect my child. In short, doing harm is compatible with acting as one ought.

I have just given a case in which one is obligated to harm. It's perhaps worth noting that in other cases, one can be obligated to help. For example, suppose I promise my spouse to do something nice for my in-laws. In such a case, I'm obligated to help my in-laws. Even if I fall a little short and thus fail to keep my promise, I may still qualify as helping my in-laws, rather than harming them.

The general lesson, then, is as follows:

> Lesson 3: The baseline should respect the fact that harm and obligation are independent.

When I say that harm and obligation are independent, I mean that it is possible to harm while fulfilling one's obligations, as in DOG ATTACK. Likewise, it is possible to violate one's obligations without harming, as in the case involving my in-laws. The baseline we propose needs to respect this fact.

This lesson should not be surprising from the perspective of morality. In acting, we often affect multiple people at once. Given this, we sometimes find ourselves in cases where in order to avoid great harm to one person, we will have to cause some lesser harm to some other person. In addition, sometimes someone morally deserves to be harmed, e.g. in the case of just punishment.[25]

---

[24]For some other related criticisms, see e.g. [Boonin, 2014, 63].

[25]Thanks to an anonymous referee for encouraging me to add this point.

Before closing this section, I want to return to the case of neglect. I said earlier that one apparent attraction of the moral minimum account is that it seems to explain why neglect is wrong. But in light of Lesson 3, there is more that can be said.

To start, let us note the following point: a powerful intuition regarding neglect is that it is morally wrong. One of the worst things that pet owners can do to their pets is to neglect them; it is totally unacceptable. But if we acknowledge that harm and obligation are independent, we can accommodate this intuition without saying that neglect is a harm. Rather, we have the following option available: we can say that neglect is very wrong, but it is not a harm.

Saying this has its attractions. The main one is that it allows us to preserve another powerful intuition, namely that other things being equal, harming is worse than failing to help.[26] For example, suppose you are tasked with watching your niece at the beach, but you ignore her and she drowns. That is obviously extremely bad. But if you had drowned her yourself, that would be even worse.

It's worth noting that I included an "other things being equal" clause in the way I stated the second intuition. Obviously, I do not wish to say that all harms are worse than all failures to help. Giving someone a minor scratch is not as bad as neglecting one's child.[27]

Again, there is surely more to be said here. But hopefully I have said enough to indicate that there might be multiple attractive positions one could take on neglect, including one on which neglect, while very bad, is not a type of harm. I think this leaves the moral minimum account in the following position: its attractions (such as its position on neglect) are somewhat weaker than they might first appear whereas its costs (such as what it has to say about Dog Attack) count very strongly against it.

5   The worse than nothing account

I am ready to introduce the hero of this story: the worse than nothing account of harm. Roughly speaking, the account says that you harm someone if you leave them worse off than if you had done nothing to them. Here is a slightly more formal statement:

---

[26]This sort of intuition is widely discussed, with some challenging it, e.g. [Rachels, 1975].
[27]Here what I say is similar to [Immerman, 2021] and [Purves, 2019, 2651]

> *The Worse Than Nothing Account Of Harm.* A's doing x harms
> B just in case it leaves B worse off than if A had done nothing to
> B.

Let me add some notes.

First, as I understand "doing x", all actions are doings. For instance, my cat's scratching me is something she does. That said there are some cases of "doing x" that are not actions. So, for instance, if a tree branch falls on me, it does something to me, viz. falling on me. But this is not a type of action. This allows us to accommodate harms that do not involve actions.

Second, the account relies on a natural contrast being "doing" and "failing to do". It only counts doings as harms, as opposed to failings to do. For instance, if you need help moving and I don't help you, I didn't do anything to you. Instead, I failed to do something, namely to help you move. Thus, according to the account, I didn't harm you, although I may have failed to help you.

Third, virtually everything one does has some sort of impact on others. So, for instance, if we're together in the same room, when I breathe I deplete the oxygen level ever-so-slightly, making it every-so-slightly harder for you to breathe. Given this, one option is to understand "doing nothing to X" to mean "doing virtually nothing to X". Another option is to understand "doing nothing to X" to mean neither harming nor benefitting X. There are various costs and benefits to these different options; for now, I shall remain neutral between them.

Fourth, regarding timing: when I talk about "doing nothing", I mean doing nothing during the same time period at which one performed one's action. So, for instance, when I ask if my son harmed his sister by knocking over her lego tower, the relevant time period is the period during which he knocked over the tower.

To further understand the worse than nothing account, it will be helpful to compare and contrast it with several related accounts.

First, there is a type of account called the "causal-counterfactual account of harm". As with all causal accounts, the causal-counterfactual account of harm say that you harm someone just in case you cause a harmful state of affairs. It differs from other causal accounts in that it says the notion of a harmful state of affairs is to be understood via a comparison with a coun-terfactual.[28] Given this, the main thing that keeps the worse than nothing

---

[28]The most well-known version of such an account is arguably due to Molly Gardner

account from counting as a causal-counterfactual account is that it asks us to focus on what one *does* and not merely what is *caused by what one does*. So, for example, if I violate your trust, the worse than nothing account will consider the violation itself in assessing whether you are worse off and not merely what the violation causes. In particular, it says that you are harmed if my violation left you worse off than if I had done nothing to you. By contrast, a causal-counterfactual account can only look at the effects of my violation and ask if these are worse than in the relevant counterfactual case. This difference seems to me to favor the worse than nothing account over the causal-counterfactual account, given that the violation can (arguably) make someone worse off even apart from its effects.

Second, there is an account I discussed earlier: the worse than the alternative account. Both the worse than the alternative account and the worse than nothing account reference a counterfactual, but their choice in counterfactuals differ.

The worse than the alternative account uses the following counterfactual: what would have happened if my action had not occurred. In possible worlds talk, this is the nearest possible world in which my action had not occurred.

By contrast, the worse than nothing account uses the following: what would have happened if I had done nothing to you. In possible worlds talk, this is the nearest possible world in which I did nothing to you.

These two worlds need not be the same. For example, suppose that you bop me on the shoulder. Suppose that if you hadn't done that, you would have pulled my hair. And suppose that if you had done nothing to me, you would have pestered the person sitting next to me. In this case, the worse than the alternative's contrast case involves you pulling my hair, whereas the worse than nothing's contrast case involves you pestering the person sitting next to me.

With the account explained, we are ready to ask: how does it do when it comes to identifying a baseline? The answer: very good indeed; the worse than nothing account accommodates each of the lessons I have discussed.

Lesson 1 was that a good baseline should hold external factors constant. The worse than nothing account does this. Its baseline alters the main event (and what counterfactually follows from it) but keeps other things constant.

[Gardner, 2015]. For a recent critique of such accounts, see [Carlson et al., Forthcoming]. As this critical article brings out, unless those offering such accounts are careful with how they frame their counterfactuals, they will run into similar problems to those facing non-causal counterfactual accounts.

Because of this, the worse than nothing account handles Slow Decline with ease. If my doctor hadn't treated me, but instead done nothing to me, I would have ended up worse off. And in such a case, the illness would have affected me even more strongly. This is why the doctor's treating me does not count as a harm.

It is worth noting that the worse than nothing account has nice things to say about the cases where causation and counterfactuals come apart. For instance, consider cases in which you preempt someone else's action. So, for instance, consider the case of Near Accident in which I push you out of the way of an oncoming car and you fall and skin your knee. Here, even though my action causes you to skin your knee, I benefit you (in an all-things-considered sense) because I preempt the actions of another – the distracted driver. The worse than nothing account has no trouble handling this case; it compares my action to a case in which I had done nothing, and the oncoming car had run into you. It concludes that my push did not harm you.[29]

This speaks to a point that I raised earlier on in this paper. I said that one seeming advantage of the temporal comparative account is that it does not rely on counterfactuals. But as we look into the details, we see that our intuitions on harm seem to track counterfactuals, as opposed to causation. So perhaps counterfactual accounts of harm are in a better place than those in other domains.

Lesson 2 was that a good baseline should be neutral. The worse than nothing account accomplishes this; opting for a descriptive understanding of neutrality. In particular, its baseline case replaces the main event with a case in which you do nothing.

Because of this, the account handles Knight with ease. If Bobby Knight had done nothing to Phil, Phil would have been better off. This is why his choking Phil counts as a harm according to the worse than nothing account.

The particular sense of neutrality that the worse than nothing account invokes has added attractions. For instance, as we have seen, it can accommodate harmful events that are not actions. Even though the tree branch falling on me is not morally assessable, we can still assess it. That is, we can

---

[29]Similar things go for cases in which I preempt my former self. So, for instance, suppose I poison your drink in an attempt to kill you. Then, thinking better of it, I knock the glass out of your hand, startling you and causing a minor injury. In this case, my pushing the glass out of your hand benefits you (in an all-things-considered sense) by preempting the actions of my former self.

still ask how I would be doing if the the tree branch had done nothing to me.

In addition, because the worse than nothing account allows us to understand neutrality in non-moral terms, it allows for the possibility of accounting for morality in terms of harm.

In offering Lesson 2, I gave a minimal understanding of neutrality, on which it involved neither harm nor help. This fits very nicely with the worse than nothing account. For instance, suppose someone complains that a police officer harmed them. If the officer responds "I didn't do anything to him" that sounds like a denial.

Compare this with what the moral minimum account would have us say. Suppose the someone complains that a police officer harmed them and the officer responds, "I was just doing what I was supposed to". This doesn't sound like a denial of harm. Rather, it sounds like an excuse – the officer seems to be giving an excuse for the harm that was incurred.

Lesson 3 was that the baseline should accommodate the independence of harm and obligation. The worse than nothing account does this by opting for a descriptive notion of neutrality. This makes room for the possibility that something might fulfill all one's obligations while still harming someone else.

Indeed, one such example is the DOG ATTACK case. Suppose that I, the parent, had done nothing to the attacking dog. In this case, the dog would have been better off; it wouldn't have been injured. Even though I was acting as I ought to have, I nonetheless left the dog worse off than if I had done nothing to it.

This allows us to take a fairly nuanced understanding of the moral significance of doing nothing. On this picture, if one does nothing, one fails to harm. And this matters, especially if we think harming is, other things being equal, less bad than failing to help. That said, doing nothing does not absolve us of moral wrongness. There may be cases, such as DOG ATTACK, that call out for us to do something. If we fail to do so – if I fail to protect my child – I will have acted wrongly.[30]

---

[30]t's perhaps worth emphasizing that a similar approach can be extended beyond cases of doing nothing. So, for example, suppose I cannot afford a very expensive treatment for my extremely sick pet, and I choose to euthanize him, rather than letting him suffer a painful death. On this approach we could say that choosing to euthanize him over the expensive treatment counts as a failure to help him, as opposed to a harm, given that if I had done nothing, he would have suffered a terrible fate. And this is perfectly consistent with the idea that in acquiring a pet, one takes on certain obligations and that a failure

I think it is no small point in favor of the worse than nothing account that it can accommodate cases that plague its rivals. But I think it's worth emphasizing that my defense of the account is not merely that it could accommodate cases. As I see things, these cases bring out general lessons regarding what we should be looking for in a baseline. Together, these lessons significantly narrow down the possibility space for accounts of this baseline.

Recall first Lesson 1, which was that a good baseline should hold external factors constant. As I noted, to do so, we seem to need to invoke a counterfactual.

Meanwhile, suppose we wish to accommodate my second two lessons, viz. that a good baseline should replace the main event with one that is neutral, and that harm and obligation are independent. Together, these lessons push us away from a moral sense of neutrality and towards a descriptive sense.

At this point, we seem to be pushed towards something very much like the worse than nothing account. In particular, in a baseline, we're looking for a counterfactual case in which I behaved in a neutral manner (in a descriptive sense) towards the recipient. This puts us in the neighborhood of the worse than nothing account.

6   Conclusion

This paper has uncovered several lessons regarding the baseline we should be providing in accounting for harm. There doubtless exist further questions to explore and lessons to discover. But hopefully what I have said has taken us some way along the direction we should be heading. And in the process I've provided some reasons to think that the worse than nothing account will prove an attractive companion in our journey.[31]

---

to live up to these can be morally wrong, even if it doesn't amount to a harm.

For another example, suppose my heart explodes, and I die as a result. On this approach, when my heart stops beating normally and instead explodes, this counts as a failure to help me (this is all assuming that my heart's exploding is no worse for me than its stopping). Again, this is not to deny that heart failure is a very bad thing. Thanks to an anonymous referee for encouraging me to discuss this sort of example.

[31]Thanks to Caleb Perl and several anonymous referees for helpful comments.

References

[Bennett, 1995] Bennett, J. (1995). *The Act Itself.* Clarendon Press.

[Boonin, 2014] Boonin, D. (2014). *The Non-Identity Problem and the Ethics of Future People.* Oxford University Press.

[Bradley, 2012] Bradley, B. (2012). Doing away with harm. *Philosophy and Phenomenological Research*, 85(2):390–412.

[Carlson et al., Forthcoming] Carlson, E., Johansson, J., and Risberg, O. (Forthcoming). Causal accounts of harming. *Pacific Philosophical Quarterly.*

[Donagan, 1977] Donagan, A. (1977). *The Theory of Morality.* The University of Chicago Press.

[Feit, 2015] Feit, N. (2015). Plural harm. *Philosophy and Phenomenological Research*, 90(2):361–388.

[Feit, 2017] Feit, N. (2017). Harming by failing to benefit. *Ethical Theory and Moral Practice*, 22(4):809–823.

[Feldman, 1991] Feldman, F. (1991). Some puzzles about the evil of death. *The Philosophical Review*, 100(2):205–227.

[Gardner, 2015] Gardner, M. (2015). A harm based solution to the non-identity problem. *Ergo*, 2:427–444.

[Hanna, 2016] Hanna, N. (2016). Harm: Omission, preemption, freedom. *Philosophy and Phenomenological Research*, 93(2):251–73.

[Hanser, 2008] Hanser, M. (2008). The metaphysics of harm. *Philosophy and Phenomenological Research*, 77(2):421–450.

[Howard-Snyder, Summer 2002] Howard-Snyder, F. (Summer 2002). Doing vs. allowing harm. *The Standford Encyclopedia of Philosophy.*

[Immerman, 2021] Immerman, D. (2021). The worse than nothing account of harm and the preemption problem. *Journal of Moral Philosophy*, 19(1):25–48.

[Johansson and Risberg, 2019] Johansson, J. and Risberg, O. (2019). The preemption problem. *Philosophical Studies*, 176:351–365.

[Kagan, 1989] Kagan, S. (1989). *The Limits of Morality*. Oxford University Press.

[Klocksiem, 2012] Klocksiem, J. (2012). A defense of the counterfactual comparative account of harm. *American Philosophical Quarterly*, 49(4):285–300.

[Norcross, 2005] Norcross, A. (2005). Harming in context. *Philosophical Studies*, 123(1-2):149–73.

[Northcott, 2015] Northcott, R. (2015). Harm and causation. *Utilitas*, 27(2):147–64.

[Purves, 2016] Purves, D. (2016). The harms of death. In Cholbi, M., editor, *Immortality and the Philosophy of Death*. Rowman and Littlefield.

[Purves, 2019] Purves, D. (2019). Harming as making worse off. *Philosophical Studies*, 176(10):2629–2656.

[Rabenberg, 2015] Rabenberg, M. (2015). Harm. *Journal of Ethics and Social Philosophy*, 8(3):1–32.

[Rachels, 1975] Rachels, J. (1975). Active and passive euthanasia. *New England Journal of Medicine*, 282:78–86.

[Roberts, 2007] Roberts, M. A. (2007). The non-identity fallacy: Harm, probability and another look at parfit's depletion example. *Utilitas*, 19(3):267–311.

[Ross, 1930] Ross, W. D. (1930). *The Right and the Good*. Clarendon Press.

[Thomson, 2011] Thomson, J. J. (2011). More on the metaphysics of harm. *Philosophy and Phenomenological Research*, 82(2):436–58.

[Woollard, 2012] Woollard, F. (2012). Have we solved the non-identity problem. *Ethical Theory and Moral Practice*, 15:677–90.