

Harming as making worse off

Duncan Purves¹

© Springer Nature B.V. 2018

Abstract A powerful argument against the counterfactual comparative account of harm is that it cannot distinguish harming from failing to benefit. In reply to this problem, I suggest a new account of harm. The account is a counterfactual comparative one, but it counts as harms only those events that make a person (rather than merely allow him to) occupy his level of well-being at the world at which the event occurs. This account distinguishes harming from failing to benefit in a way that accommodates our intuitions about the standard problem cases. In laying the groundwork for this account, I also demonstrate that rival accounts of harm are able to distinguish harming from failing to benefit only if, and because, they also appeal to the distinction between making upshots happen and allowing upshots to happen. One important implication of my discussion is that preserving the moral asymmetry between harming and failing to benefit requires a commitment to the existence of a metaphysical and moral distinction between making and allowing.

Keywords Harm · Benefit · Doing/allowing distinction · Counterfactual comparative account of harm · Acts · Omissions

Moral common sense holds that harming a person is a graver offense than merely failing to benefit him, even if each form of treatment leaves him in the same position with respect to his well-being. If I refrain from improving the quality of toxic drinking water in a developing country, and several people die as a result, I fail to benefit those people. If I poison the country's water supply, and several people die

✉ Duncan Purves
duncan.purves@gmail.com

¹ Department of Philosophy, University of Florida, 330 Griffin-Floyd Hall, Gainesville, FL 32611-8545, USA

as a result of the poisoning, I harm them. The latter form of treatment, it is commonly thought, is morally much worse than the former *because* poisoning someone is a harm for that person. To justify harming a person, I must show that some number of conditions have been satisfied: that the victim deserves it; that harming him is necessary to prevent greater (perhaps much greater) harm to others; that I have some special right to harm him; or that he has consented to my harming him.¹ There is no comparable need to satisfy these conditions in order to justify failing to benefit someone. Describing this moral asymmetry, Seana Shiffrin writes.

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 (Shiffrin 2012: 361).

A plausible theory of harm (or harming)² must account for the intuitive distinction between harms and failures to benefit and the moral asymmetries that this distinction underpins.

One powerful objection to the counterfactual comparative account of harm is that it cannot respect the moral and metaphysical distinction between harms and failures to benefit.³ Consider one formulation of counterfactual comparative account of harm:

CCA An event *e* is a harm for a subject S just in case S is better off in the nearest possible world in which *e* does not occur than S is in the relevant e-world.⁴

CCA is an account of overall, all-things-considered harm. It is an account of all-things-considered harm, as opposed to *pro tanto* harm, because, when determining whether some event is a harm, it takes into account all of the ways a person is better (or worse) off than he would have been had the event not occurred. CCA is an account of the *overall* harmfulness of an event in that it considers both the event's intrinsic and extrinsic value in determining whether the event is a harm. Some harms are intrinsically bad. Suffering agony is a harm, and it is intrinsically bad for the person who suffers it. Some harms are only extrinsically bad. Being prevented from receiving life-saving surgery is a harm, but it is not intrinsically bad; it is harmful because it deprives a person of intrinsic goods.

¹ Shiffrin (2012: 362).

² More on the difference between harm and harming below.

³ Seana Shiffrin (1999, 2012), Ben Bradley (2012: 397), and Matthew Hanser (2008). Klocksiem (2012) introduces the objection as well.

⁴ This formulation of the counterfactual comparative account is Klocksiem's (2012). Hanna (2015) formulates the view in a similar way. I follow them in using the possible worlds analysis of counterfactuals for the sake of clarity (Stalnaker 1968). Doing so facilitates explanation of the problems for the counterfactual comparative account of harm. I am not assuming realism about possible worlds.

CCA is an attractive account of harm for several reasons. First, it obeys *ontological* and *normative neutrality* in that it does not restrict harmful events to actions, and it does not entail that all harming is wrongful.⁵ Natural events might be harms for someone. A tree's falling onto me, breaking my leg, harms me, but this event is not a wrongful harm. Second, CCA makes harm something we have reasons to care about. It seems that I always have prudential reason to care about being better rather than worse off, and, if the occurrence of some event makes *no* difference to how well off I am, it is not clear how it could be a harm for me. Third, CCA can explain the harmfulness of deprivations, including the deprivation suffered in death, which other accounts struggle to do.⁶

CCA can also be used to analyze the conditions under which an agent counts as harming a victim. Consider a counterfactual comparative account of agential harm:

CCAH An agent X harms a victim S, *qua* agent, just in case X performs an action A such that S is better off in the nearest possible world in which A does not occur than S is in the relevant A-world.⁷

The qualification '*qua* agent' is included in CCAH to distinguish distinctly agential harm from ways that an agent might harm someone other than by acting. I might harm someone by falling onto him after being pushed.⁸ Though falling onto someone is a harmful event in which I figure, I do not harm someone by *acting* in these cases. I figure in the harmful event as an instrument, the same way a gun or a rock might figure in a harmful event. Agents may be able to harm individuals without acting by figuring in harmful events the way that mere objects figure in harmful events, but I am not interested in an analysis of that sort of harming.

The trouble for CCA is that it counts some events as harms that are, intuitively, mere failures to benefit. Several illustrations of the problem can be found in the literature.

Robins' Clubs Suppose that Batman purchases golf clubs with the intention of giving them to Robin, but the Joker persuades him to keep them for himself. Had Batman not kept the clubs he would have given them to Robin. (Bradley 2012: 397)

⁵ Not all analyses of harm have this virtue. See, e.g. Roberts (1998: 32).

⁶ Bradley (2012) proposes these criteria for a successful account of harm. See Purves (2015, 2016) for an extended discussion of the unique ability of CCA to handle the harm of death. Bradley (2012) and Feit (2016) also discuss CCA's treatment of death.

⁷ Thus the account of agential harm is quite different from Feinberg's (1986) moralized account of agential harm. In part, this difference is due to Feinberg's legal interest in using the concept of harm to determine when compensation is owed to a victim. Insofar as the counterfactual comparative account presented here diverges from Feinberg's, I take this to be a virtue. It seems that one person can harm another without acting wrongly, and without owing the victim compensation, contrary to Feinberg's account.

⁸ I thank Neil Feit for suggesting these examples to me. See Feit (2017) for other examples in which agents may be said to harm without acting.

Jughead's Dollar Archie has agreed to lend Jughead one hundred dollars, and...as Archie is handing the money over, at the last second, he notices a one-dollar bill sticking to the twenties, unsticks it, and puts it back in his wallet. Although Jughead would be better off if Archie had left the one in with the twenties, this does not appear to rise to the level of genuine harm. (Klocksiam 2012: 294).

The actions performed in Robin's Clubs and Jughead's Dollar are thought to be morally different from the actions performed in cases of preventative harming.

Joker's Removal Batman has delivered golf clubs to Robin, and the Joker removes the clubs just before Robin opens the door to retrieve them. Had the Joker not removed the clubs from Robin's porch, Robin would have found them and thereby been better off.

CCA classifies all of these actions as harms. CCA entails, correctly, that the Joker's removing the clubs is a harm for Robin. But CCA also entails that Batman's keeping the clubs is a harm for Robin, because Robin is better off in the nearest possible world in which Batman does not keep them. It also entails that Archie unsticking the dollar harms Jughead. But these do not seem to be harms. Batman does not in any way harm Robin by keeping the clubs. Nor does Archie harm Jughead. CCA thus misclassifies some mere failures to benefit as harms. This misclassification poses two significant problems for CCA. First, because it counts too many events as harms, CCA is *extensionally inadequate*. As Ben Bradley puts the point, the true account of harm must "fit the data."⁹ If an account of harm fails to accommodate our firmest intuitions about cases of harm, this is a strike against the theory. The second problem is that, as I described above, ordinary morality includes a commitment to the moral asymmetry between harms and benefits: harming someone is a much graver moral offense than failing to benefit him. CCA's misclassification of some failures to benefit as harms calls into question its ability to preserve the moral asymmetry between harms and failures to benefit. The Joker does something wrong. Batman does not. A straightforward explanation of this fact is that only one of their actions is a harm for Robin. But this straightforward explanation is unavailable to the proponent of CCA; according to CCA, the two actions are morally on a par with respect to their status as harms.

In this paper, I propose that a proponent of CCA must draw on a feature of rival accounts of harm in order to satisfy ordinary intuitions about the distinction between harms and mere failures to benefit. A number of rival accounts, including temporally comparative accounts (Perry 2003), event-based accounts (Matthew Hanser 2008), and non-comparative state-based accounts of harm (Seana Shiffrin 1999, 2012; Elizabeth Harman 2004) avoid the problem of distinguishing harms from failures to benefit, because (a) they draw a sharp distinction between *harms* and *harmings* and (b) they characterize *harmings* as those events which make (rather than merely allow) a person to suffer a *harm*. Drawing on these features of

⁹ Bradley (2012: 394). It is important to acknowledge that, taking all cases into account about which we have firm intuitions, CCA may do a better job fitting the data than alternative theories.

rival accounts, I revise CCA. The result is a counterfactual comparative account of harm that can distinguish, in an intuitively satisfying way, between harms and mere failures to benefit.¹⁰ According to this account, a harm is an event that makes a person—rather than merely allows him to—occupy his level of well-being at the world at which the event occurs *and* is such that the person is worse off than he would have been had the event not occurred.

My discussion yields two important upshots: First, if CCA is to “fit the data,” it must come to more closely resemble its rivals by making use of the making/allowing distinction; second, the purported metaphysical and moral distinctions between harming and failing to benefit depend on the metaphysical and moral distinctions between making and allowing. If the distinction between making and allowing is untenable, then so is the distinction between harming and failing to benefit.

In Sect. 1 I critically evaluate three recent replies on behalf of CCA to the problem of distinguishing harms from failures to benefit. In Sect. 2 I highlight the ways in which the making/allowing distinction figures in rival accounts, and I explain how appealing to the making/allowing distinction enables these accounts to distinguish events that harm from those that merely fail to benefit. In Sect. 3, I draw on the observations in Sect. 2 to argue that a modified version of the counterfactual comparative account of harm, one that appeals to the making/allowing distinction, offers a promising solution to the problem of distinguishing harms from failures to benefit. In Sects. 3 and 4 I address objections to this modified account. Section 5 concludes with a brief reflection about the importance of my arguments for moral theorizing.

1 Hanser, Kloksiem, and Hanna on harms and failures to benefit

In this section I canvas three recent replies to the problem of distinguishing harms from failures to benefit. Matthew Hanser argues that CCA does not mistake failures to benefit for harms, because CCA is an account of harmful *events*, failures to benefit are omissions, and omissions are not events. Batman’s keeping the golf clubs is not a harm because Batman’s keeping the clubs (i.e. *not* giving the clubs to Robin) is not an event; it is a non-occurrence or an omission. When Batman keeps the clubs, no event occurs such that had the event not occurred, Robin would have been better off.¹¹ Because CCA applies only to events, in cases where a person is worse off than he otherwise would have been had some event not failed to occur, CCA does not classify anything as a harm for that person. Batman’s failing to give the clubs to Robin is one such case.

Hanser’s reply shows that CCA does not entail that an agent harms a person *in every case where* he fails to benefit that person. Hanser’s reply does not, however, constitute a complete solution to the problem of distinguishing harms from failures to benefit. Let us grant that omissions are not events and that failures to benefit are

¹⁰ I will not address other problems for the counterfactual comparative account, including cases of preemption and overdetermination and the non-identity problem. Hanser (2008) and Bradley (2012) discuss these problems in some detail, and Feit (2015) and Kloksiem (2012) offer some solutions.

¹¹ Hanser (2008: 427).

omissions. Still, in some cases where a person merely fails to receive a benefit, there is an event in the vicinity for which it is true that, had that event not occurred, the person would have been better off. Consider a version of *Robin's Clubs* with some details filled in.

Robin's Clubs⁺ Batman purchases golf clubs with the intention of giving them to Robin. As Batman is loading the clubs into the trunk of the Batmobile, it occurs to him that he hasn't checked the clubs for defects. Batman takes a few swings with the clubs. Because the feel of the clubs is much better than his own clubs, Batman decides to keep them. He tucks them away in his garage.¹²

Here Batman performs an action—swinging the clubs—such that, had he not performed that action, he would have given the clubs to Robin. Archie's unsticking the one-dollar bill is an event such that, had Archie not performed it, Jughead would be wealthier. Whether or not omissions are events, CCA entails that Batman swinging the clubs is a harm for Robin and that unsticking the dollar is a harm for Jughead. This seems incorrect.¹³ So, Hanser's appeal to the distinction between events and omissions shrinks the number of cases in which CCA misclassifies failures to benefit as harms, and it thereby succeeds in showing that CCA does not entail that an agent harms someone *whenever* he fails to benefit that person. Still, CCA will sometimes and counterintuitively count as harms the actions that agents perform instead of benefiting others.

Justin Klocksiem (2012) argues that the difference between harming and failing to benefit can be explained by appealing to the effect of contextual features of cases on our judgments about whether the closest possible world is one in which the subject is better off. One contextual detail that has this effect is whether the agent is passive or active. Klocksiem says, "In most contexts, the most salient comparison case is one in which the relevant agent does what would intuitively be thought of as refraining from action." (Klocksiem 2012: 294). According to Klocksiem, our judgment that Archie does not harm Jughead by unsticking the dollar can be explained by the fact that our default comparison in this case is *not* with a world in which Archie gives Jughead the twenties and leaves the extra dollar. Our default comparison is with the world in which Archie refrains from action altogether. In that world, Jughead is *not* better off, because, in that world, Archie doesn't give him the twenties. Klocksiem also says that whether an agent is morally obligated to act in a certain way can affect our intuitions about what would have happened had he acted differently. He asks us to suppose that Betty is ill and that Archie has promised to serve as her nurse, but then he spends the day with Veronica instead, and Betty does

¹² This revised version of *Robin's Clubs* is inspired by Feit (2017) who offers examples in which it appears that the agent *actively* fails to benefit a person.

¹³ If we accept Neil Feit's 'plural' version of the counterfactual comparative account of harm, we may be committed to the result that Robin suffers a harm even in cases where Batman fails to give Robin the golf clubs in the nearest possible world in which he does something other than swing the clubs. This will happen any time when swinging the clubs is a member of the smallest super-plurality of every plurality of events P such that (1) if none of the events in P had occurred, Robin would have been better off; and (2) there is no smaller sub-plurality of P such that if none of the events in it had occurred, Robin would have been better off. (Feit 2015: 17).

not get better. Klocksziem says that serving as Betty's nurse is the default action, because Archie is obligated to do so. Thus, the salient possible world in which Archie does not spend the day with Veronica is one in which Betty is better off, because, in that world, Archie tends to her. It thus seems that Archie harms Betty.¹⁴ These two factors together can explain the intuitive harmfulness of Joker's stealing the clubs: Joker is active in the example, and he does the wrong thing. We tend to judge that stealing the clubs from Robin is a harm for him, because our default comparison is with a world in which the Joker does the *right* thing by *refraining* from stealing the clubs—that world seems 'closest' given the context.

Klocksziem's explanation works in some cases, but it is hard to see how it works when the counterfactual details are made clear. When the counterfactual details are clearly described, we should have the intuition that failures to benefit are harms, if CCA is true. But this is not what happens. For instance, it is explicit in Robin's Clubs⁺ that, had Batman not swung the clubs, *he would have* given them to Robin. Though Batman has not made any *promises* to Robin to do this, it should nonetheless be clear (because it is explicit) that the nearest possible world in which Batman does not swing the clubs is one in which Robin is better off. Still, most people do not have the intuition that Batman's action is a harm for Robin. It is not obvious how the fact that Batman is passive and acts permissibly in the example is supposed to override the explicit details concerning what would have happened had Batman not taken a few swings with the clubs.¹⁵

Nathan Hanna (2015) argues that CCA gets things exactly right: Batman keeping the golf clubs harms Robin. Suppose that the Joker persuades Batman to keep the golf clubs. This, Hanna thinks, is a harm for Robin. Hanna argues that it follows from this that Batman harms Robin:

- 1) If Batman's [keeping the clubs] doesn't harm Robin, the Joker's actions don't harm Robin.
- 2) The Joker's actions harm Robin.
- 3) So, M is true: Batman's [keeping the clubs] harms Robin.¹⁶

Hanna's response is persuasive only if we are more confident that the Joker harms Robin by persuading Batman to keep the clubs than we are that Batman does not harm Robin by keeping the clubs. We could just as easily substitute premise (2) with the following premise:

- 2*) Batman's actions do not harm Robin.

¹⁴ Klocksziem (2012: 294).

¹⁵ Klocksziem is here attempting to explain why our intuitions go awry, because he denies that there is any genuine metaphysical distinction between harming and failing to benefit.

¹⁶ Hanna (2015: 3). I have substituted the bracketed text for Hanna's phrase "changing his mind" to maintain consistency with the Bradley's original description of the example.

This would yield:

3*) So, the Joker's actions do not harm Robin.

It is hard to see why we should be more confident that (2) is true than that (2*) is true, and Hanna does not provide an argument for thinking that (2) is more plausible than (2*). It would therefore be preferable to respond to the problem without relying on Hanna's argument.¹⁷

2 The role of the making/allowing distinction in competing analyses of harm

Existing responses to the problem of distinguishing harms from failures to benefit are unsuccessful. To address the problem, we must amend CCA. To set the stage for my proposed amendment to CCA in Sect. 3, in this section I show that several of the most prominent rivals to CCA invoke the making/allowing distinction, and that this feature of the accounts allows them to distinguish harming from failing to benefit. The takeaway lesson is that CCA must also look to the making/allowing distinction to distinguish harms from failures to benefit.

To avoid some confusion in the discussion that follows, I want to first highlight a feature absent from CCA that is shared by its rivals. The rivals to CCA typically draw a sharp distinction between *harms*, on the one hand, and events that *harm* (or are *harmings of*) a person on the other. On these views, harms are (typically) identified with states that it is in some way bad for a person to be in, or with events that it is bad for a person to have happen to him. *Harming* is an active notion. An event *harms* a person (or is a *harming of* a person) if and only if it causes—or, as I will soon suggest, makes—that person to suffer a harm. CCA is distinctive in part because it only provides conditions under which an event counts as a harm.

2.1 Temporally comparative account

Stephen Perry defends a temporally comparative account of harm. He calls this account the “historical worsening condition” (HWC).

HWC A person has been harmed only if some relevant interest of that person has been affected adversely, meaning the interest has been caused to worsen or deteriorate in time.¹⁸

If we assume that a person has *been* harmed only if some event has *harmed* him, then HWC can be understood as offering an analysis of the conditions under which an event harms (or is a harming of) a person. An event harms a person only if it causes one of his interests to worsen in time. But what of *harm*? Because Perry does

¹⁷ Like Nathan Hanna, Feit (2017) argues, more convincingly to my mind, that we should count failures to benefit as harms. I do not here have space to address Feit's many interesting arguments.

¹⁸ Perry (2003: 1295).

not offer an account harm, but only harming, I will offer a charitable suggestion. The following principle seems to be in the spirit of Perry's view.

An event e constitutes a harm for a subject S if and only if e is a decline in S 's well-being.

Despite the appearance of the word 'cause' in HWC, I think we should understand Perry's view as invoking the making/allowing distinction rather than the cause/mere condition distinction. This is because some events are harms for persons in virtue of their non-causal consequences. For example, if some version of the desire satisfaction theory of well-being is true, then events can make a person's well-being decline non-causally. Suppose that my mother desires that I not join a cult. If I join a cult, this frustrates her desire, making her occupy a lower well-being level than she did before I joined the cult. But, it would misconstrue things to say that my joining a cult *causes* her desire to be frustrated. The event makes her well-being decline by bringing to completion an event that is bad for her: *her desiring that I not join a cult and my joining a cult*.¹⁹ My joining a cult harms my mother by *constituting* her desire frustration. The desire satisfaction theory of well-being is controversial, but a general theory of harm or harming should not rule out the possibility that a person could be harmed in the way described above. To accommodate events that harm persons non-causally, HWC must be revised to say that events harm by *making* some relevant interest of S worsen or deteriorate in time.²⁰

So, on the best version of Perry's view, whether an event e harms S depends on whether e makes S 's interest worsen or deteriorate in time. This feature of the account is crucial if HWC is to distinguish between harming and failing to prevent harm. Imagine that Steve will be struck by an oncoming bus unless I push him out of the way. I choose instead to call for help. After I call for help, Steve is hit by the bus and maimed. Steve is worse off after I call for help than he was before I called for help. HWC says that this decline in well-being is clearly a *harm* for Steve. If Perry is to avoid the implication that I harm Steve by calling for help, he must maintain

¹⁹ Fiona Woollard (2008: 272–273), following Jonathan Bennett (1995: 38), provides additional examples illustrating how the difference between making and allowing can be applied to non-causal upshots.

²⁰ Of course, Perry, as well the theorists I consider later, may be invoking the distinct but related concept of causation. It is not immediately clear, however, how invoking causation would help these theorists distinguish harms from mere failures to benefit. For example, Lewis's (1973) counterfactual analysis of causation would clearly not help, and the authors I consider here do not provide detailed analyses of causation. An account of causation that requires the transfer of energy will be of no help in explaining how removing the net from the falling acrobat counts as a harm. Mackie's 'INUS condition' for causation can explain the harmfulness of removing the net, because, on this account, removing the net is a cause of the death of the acrobat (i.e., removing the net is an insufficient but necessary component of a collection of factors that are themselves unnecessary but sufficient for the death) (Mackie 1965). But Mackie's account, combined with the temporally comparative account, also implausibly entails that withdrawing a barrier is, in general, a harm, even when sustaining the barrier requires the use of resources belonging to the agent who withdraws the barrier. I am also unaware of an account of causation that would maintain a distinction between the case in which the Joker removes the clubs from Robin's porch and the case in which Batman returns the clubs to his garage from his own stoop.

that my calling for help does not make, but merely allows, some relevant interest of Steve's to deteriorate in time.

2.2 Event-based account

Matthew Hanser's event-based account of harm analyzes harm in terms of losses of basic goods. Basic goods, on this view, are not states that it is good (or better) to be in. Rather, they are things that it is "good to have."²¹ He says, "the basic goods are those the possession of which makes possible the achievement of a wide variety of the potential components of a reasonably happy life."²² Hanser provisionally includes physical powers and capabilities in the category of basic goods. The power of sight is one example. Hanser develops his account of *harm* using the notions of losing and gaining quantities of basic goods. Someone suffers (enjoys) a 'non-derivative' harm (benefit) with respect to some basic good if and only if he loses (gains) some quantity of that good. When someone loses his sight, he suffers a non-derivative harm. On Hanser's view, someone suffers a 'derivative' harm (benefit) if and only if he is prevented from enjoying (suffering) a non-derivative benefit (harm) with respect to a basic good.²³ Of *harming* Hanser says, "I shall assume, provisionally, that being harmed is roughly equivalent to being caused to suffer it."²⁴ Because the example does not involve the loss of any basic goods, I am unsure whether Hanser's account is able to accommodate the intuition that my joining a cult is a harm for my mother. Nonetheless, if it is to have any hope of doing so, Hanser, like Perry, must replace language of 'causing' with language of 'making'. When we do this, Hanser's view is that a subject's being harmed is equivalent to being made to suffer it. From this I think we can infer that *harming*, on Hanser's view, is to make a subject suffer harm. This at least leaves open the possibility that I harm my mother by joining a cult.

Like Perry, Hanser needs to restrict events that harm to those that *make* a subject suffer harm in order to explain why events that fail to prevent harm to a person do not themselves harm that person. The event-based analysis of suffering harm correctly entails that Steve suffers harm when he is struck by the bus, because Steve loses some quantity of basic goods. Because it makes Steve suffer this harm, the bus striking Steve harms him. But why doesn't my calling for help harm Steve? The answer must be that my calling for help does not make, but merely allows, Steve to suffer a harm.

2.3 State-based account

Seana Shiffrin identifies harms with multifarious harmed conditions: physical injuries, many physical disabilities, many mental disabilities, some material

²¹ Hanser (2008: 440).

²² Ibid: 440.

²³ Ibid: 442.

²⁴ Hanser (2008: 421).

inabilities, incidents of pain, the failure or ruin of certain sorts of important projects and relationships, some losses, and death. On her view, conditions like disability, injuries, and illnesses are typically harms because they “often significantly impede one’s capacity to achieve substantial congruity between one’s will and one’s life.”²⁵ So, what each harmed condition has in common is that each is a state such that a person in that state would rationally will not to be in it. Bradley restates Shiffrin’s view so that *harming* is analyzed in terms of what one rationally wills.

An event harms someone iff it causes him to be in a state that he rationally wills not to be in.²⁶

The term “causes” only appears in Bradley’s summary of Shiffrin’s account, but Shiffrin has indicated that she would accept Bradley’s characterization.²⁷ Though the role of the making/allowing distinction is downplayed in Shiffrin’s work, as we saw with Hanser’s and Perry’s accounts, the distinction is crucial if her account is to handle failures to prevent harm. I have the opportunity to prevent the bus from striking Steve, but instead I call for help. Soon thereafter, Steve is in a harmed state, one that he rationally wills not to be in. Shiffrin’s account must explain why it is that *I* do not harm Steve by calling for help, whereas I would have harmed Steve had I *pushed* him in front of the bus. The explanation must be that in calling for help I merely allow Steve to be in a state in which he rationally wills himself not to be.²⁸

The purpose of this section was to show that each account of harm that successfully distinguishes harming from failing to benefit (by failing to prevent harm) does so only because it invokes the making/allowing distinction. I propose that this is no accident. There is a lesson to be learned here by proponents of CCA: if some version of CCA is to fit the data and respect the intuitive moral asymmetry between harms and failures to benefit, it will need to make use of the making/allowing distinction. Of course, unlike its rival accounts, CCA cannot do this at the level of its analysis of *harming*, because it doesn’t offer one. Still, in the next section I will show that CCA can make use of the making/allowing distinction in its analysis of harm.

3 Harming as making worse off

In this section and the next I develop an account of harm that respects our intuitions about harming and failing to benefit. The account has a counterfactual comparative element, but it restricts harms to those events that make someone occupy his level of well-being at the world at which the event occurs. The metaphysical ground of the distinction between making and allowing is notoriously difficult to pin down, and its

²⁵ Shiffrin (2012: 386).

²⁶ Bradley (2012: 400).

²⁷ Shiffrin (2012: 368, fn. 24).

²⁸ I will not discuss Elizabeth Harman’s account of harm, because it is sufficiently similar to Shiffrin’s account, and space is limited (Harman 2004).

moral significance is a topic of lively ongoing debate.²⁹ Though I will offer a limited defense of my proposed account of harm, I do not intend to settle the larger debate about making and allowing. My aim is to understand *how* to retain an account of harm in which counterfactual comparisons play a central role while also doing justice to ordinary intuitions about the distinction between harms and failures to benefit.

To begin I provisionally adopt Fiona Woollard's (2008; 2015) account of the making/allowing distinction, which is inspired by Philippa Foot's influential account.³⁰ I adopt Woollard's account because her analysis and defense of the making/allowing distinction are the most rigorous of which I am aware. Because her view draws on features of Foot's famous account, its starting point will also be familiar to many readers.

Woollard is primarily concerned with offering conditions under which an agent counts as doing harm. On her analysis, an agent A does harm if and only if some fact about A's behavior is part of a sequence leading to a *harmful* upshot. Because I am here concerned with developing an analysis of harm, and because Woollard does not present her own analysis of harm, Woollard's focus on *harmful* upshots threatens to introduce unnecessary complications to my project here. These complications can be avoided, however, by recognizing that Woollard's analysis can be applied to *any* upshot. Just as an agent can make *harmful* upshots obtain, an agent can make non-harmful upshots obtain. I will therefore present Woollard's analysis in terms of *upshots* understood broadly to include harmful and non-harmful upshots. With this in mind, the starting point of Woollard's analysis is the following:

An agent X makes some upshot U obtain if and only if some fact about X's behavior ϕ is part of a sequence leading to U.

There is a great deal to unpack here. To begin, we might ask, What are the conditions under which a fact about X's behavior ϕ counts as *part* of a sequence leading to U? On Woollard's analysis, a fact about ϕ is a part of a sequence leading to U if and only if that fact is relevant to U via an unbroken chain of *substantial* facts. If the fact about ϕ is relevant to U through any non-substantial facts, then it is not part of the sequence leading to U.³¹ Woollard identifies several ways for a fact to count as substantial. A fact is *purely substantial* if and only if (i) it is a *specificity positive* or *scalar positive* fact about A's behavior or (ii) the fact violates our *normal presuppositions*. Again, there is a great deal to unpack. I will set aside the concept of scalar positivity, because it will not play a role in my application of Woollard's analysis to a theory of harm.

²⁹ Key figures in this debate include Rachels (1975), Foot (1967, 1977), Thomson (1976), Bennett (1995).

³⁰ Foot (1967).

³¹ Woollard (2015: 29–36). Woollard leaves the concept of relevance unanalyzed except to say that a fact's being relevant to an upshot requires that the upshot in some way *depends on* the fact.

A fact about A's behavior is specificity positive if and only if it is a fact about what *is* the case rather than a fact about what is not the case.³² For example, my rolling the bowling ball down the lane is relevant to the falling of the pins through a chain of purely positive facts. The fact that I rolled the ball is a positive fact; the fact that the ball continued on a straight line toward the pins is a positive fact; the fact that the ball struck the pins is a positive fact. My behavior is therefore part of the sequence leading to the falling of the pins. Therefore, I count as *making* the pins fall. Facts about my bowling partner's behavior are also relevant to the pins falling, but at least some of the facts through which any fact about his behavior is relevant to the pins falling are negative (e.g., that he did *not* trip me during my approach), so he counts merely as allowing the pins to fall.

But this story is too simple. In some cases, a fact about an agent's behavior counts as part of the sequence leading to an upshot, even though it is relevant to the upshot through a chain of facts some of which are negative. This is because negative facts can sometimes be substantial. One way for a negative fact to be substantial is for the negative fact to violate normal presuppositions. Consider Woollard's remarks about cases where negative facts violate normal presuppositions.

We take it for granted that the air will contain oxygen. So the negative fact that there is no oxygen present contradicts the normal presuppositions. When a fact contradicts the normal presuppositions, it is substantial. Such facts conflict with our background assumptions about the world. They are highly unexpected. This means that they do not count as mere background conditions whether they are positive or negative. Thus if Johnny asphyxiates because I use a machine that sucks the oxygen from the room, the fact that there was no oxygen in the air is part of the sequence leading to Johnny's death (Woollard 2015: 58).

To take a less morally-laden example, suppose that my neighbor's houseplants are thriving because there is *no longer* shade obstructing the sun's rays, after I generously removed a large tree from his yard. Even though my behavior is relevant to the thriving of the plants through a negative fact—the fact that the tree in my neighbor's yard is absent—on Woollard's view, my behavior counts as a part of the sequence leading to the thriving of the plants, because the negative fact violates normal presuppositions. Normally, the absence of foliage on a person's property is not due to the actions of their *neighbor*!³³

Violating normal presuppositions is one way for a negative fact to count as substantial. Explaining the second way that a negative fact can be a substantial fact requires discussion of a special complication for the making/allowing distinction: events

³² Ibid: 40–53. Woollard offers an extended analysis of the conditions under which a fact counts as a specificity positive fact—that is, a fact about what is (rather than what is *not*) the case. Drawing on Jonathan Bennett's (1995) influential work, Woollard proposes that a fact about A's behavior is specificity positive, just in case most of the ways A could have moved would have failed to make the corresponding proposition about A's behavior true. I leave out this extended discussion for the sake of space.

³³ Ibid: 57–59.

that involve removing a barrier to an upshot. Sometimes removals of barriers seem to make an upshot happen. Other times they seem merely to allow an upshot to happen. Compare:

Drive Away 1 Bob's car is in the path of a boulder headed toward Victor. If Bob leaves his car where it is, it will bring the boulder to a halt. Bob needs the car to drive to the hospital for treatment for a snakebite. He drives his car out of the path of the boulder. The boulder strikes and kills Victor.

Drive Away 2 All details are the same as Drive Away 1 except that Victor's car is in the path of a boulder headed toward Victor, and Bob drives Victor's car to the hospital.

Drive Away 3 All details are the same as Drive Away 1 except that a third party Phil's car is in the path of the boulder headed toward Victor, and Phil put the car there because he wanted to prevent Victor's death. Bob drives Phil's car to the hospital.³⁴

Concerning Drive Away 1, Woollard says that it is clear that Bob allows Victor's death. In Drive Away 2 and 3, it seems that Bob makes Victor's death occur. To deal with these cases, Woollard introduces the notion of *relatively substantial facts*. Starting with Jeff McMahan's influential (1993) analysis of the conditions under which the removal of a barrier counts as making an upshot obtain, Woollard settles on the following analysis:

A fact is *relatively substantial* if and only if it is a negative fact about the absence of a barrier to U, where either (a) sustaining the barrier requires the use of resources belonging to someone other than the person who removed the barrier who does not want U to obtain, or (b) someone other than the person who removed the barrier has a non-need-based claim to the barrier to U.³⁵

I will focus on condition (a) and set aside condition (b), because (b) is not important for what follows. By condition (a) of Woollard's analysis, in Drive Away 1, Bob's driving the car is not part of the sequence leading to Victor's death, because the fact that Bob drove the car out of the path of the boulder is relevant to the upshot only through a negative fact about the absence of the car, and the sustenance of the barrier imposed by the car requires the continued use of resources belonging only to *Bob*. Therefore, Bob counts as merely allowing Victor's death. In Drive Away 2, sustaining the barrier to Victor's death requires the continued use of Victor's resources (we can assume further that Victor does not want to die). Therefore, a fact about Bob's behavior is relevant to Victor's death through a *substantial* negative fact about the absence of the car, and so the fact about Bob's behavior *is* part of the sequence leading to Victor's death. Therefore, Bob counts as *making* Victor's death

³⁴ These cases are paraphrased from Woollard (2015: 62).

³⁵ Ibid: 69–78. The situation becomes more complicated if, say, the Bob's poisoning is due to actions on the part of the third-party owner of the car, the same car that now stands as a barrier against the threat posed to Victor. I set this complication aside (74).

obtain. In Drive Away 3, again, sustaining the barrier to Victor's death requires only the continued use of resources belonging to someone other than Bob (who also does not want Victor to die).³⁶ To sum up the features of Woollard's account that are relevant to my discussion: an agent counts as making an upshot happen if and only if a fact about his behavior is part of a sequence leading to the upshot; a fact about an agent's behavior is part of a sequence leading to an upshot if and only if it is relevant to the upshot through an unbroken chain of substantial facts; substantial facts are positive facts, negative facts that violate normal presuppositions, and negative facts about the absence of barriers where the resources required to sustain the barrier belong to someone other than the agent (and where the person to whom the resources belong does not want the upshot to happen).

Below, I will use Woollard's analysis of the making/allowing distinction to showcase a new counterfactual comparative account of harm. But a successful account of harm needs to be ontologically neutral—its application should not be limited to actions—and Woollard's account only provides conditions under which an *agent* counts as making an upshot occur. So, some modification is in order. On Woollard's account, an agent counts as making an upshot occur if and only if one of the agent's actions is part of a sequence leading to the upshot. But actions are simply a species of events, and events that are not actions can be parts of sequences leading to upshots. Even if a tree's falling alone in the woods does not make a sound, it can make *something* happen. In principle, then, Woollard's analysis of the making/allowing distinction can be modified to apply to events that aren't actions. I propose the following modification.

Making and Allowing for Events

- (i) An event e makes U occur if and only if e is part of the sequence leading to U , where e is part of the sequence leading to U if and only if a fact about e is relevant to U through a complete chain of substantial facts.
- (ii) e merely allows U to occur if and only if e is relevant to U but not part of the sequence leading to U .

By paying attention to the distinction between events that make upshots happen and those that merely allow upshots to happen we can formulate an account of harm that retains a counterfactual element while respecting the difference between harms and failures to benefit. I will now describe one such view. I call it 'Harming as Making' (HAM):

HAM An event e is a harm (benefit) for S if and only if (1) e makes S occupy S 's well-being level in the e -world, and (2) S 's well-being level is higher (lower) in the nearest world in which e does not occur. An event e is a failure to benefit (harm) S if and only if (3) e does not make S occupy S 's

³⁶ In Drive Away 1–3, the car belongs to the person it does because that person *owns* it. But, Woollard claims, belonging is a weaker notion than ownership. An apartment can belong to the person renting it, but the renter does not own the apartment. I assume that Woollard's verdicts about Drive Away 1–3 would remain the same if the car was stipulated to be a rental.

well-being level in the e -world, and (4) S 's well-being level is higher (lower) in the nearest world in which e does not occur.³⁷

HAM has the features that make CCA an attractive account of harm: it accounts for the harm of death (more on this below); it allows for a symmetrical treatment of benefits and harms; it is ontologically neutral because it leaves open the possibility that natural events and actions performed by non-agents can be harms or failures to benefit; it makes harms something we have reason to care about; it explains what all harms and failures to benefit have in common; finally, it lends itself to a plausible account of agential harm and failure to benefit:

AHAM An agent X harms (benefits) a subject S , *qua* agent, just in case there is some event e such that e is a harm (benefit) for S , and e is an action (behavior) performed by X . Otherwise X merely fails to benefit (harm) S

For my purposes, the most attractive feature of HAM is that it accounts for the distinction between harming and failing to benefit.³⁸ A fourth variation of Woollard's Drive Away illustrates this feature.

Drive Away 4 James dislodges a boulder from a hillside, sending it careening downward toward Victor. Bob's car is in the path of a boulder headed toward Victor. If Bob leaves his car where it is, it will bring the boulder to a halt. Bob needs the car to drive to the hospital for treatment for a snakebite. He drives his car out of the path of the boulder. The boulder strikes and maims Victor.

HAM counts James's dislodging the boulder as a harm for Victor because (1) dislodging the boulder makes Victor occupy his well-being level at the world at

³⁷ Notice that this account is not the same as the 'causal comparative analysis' discussed by Bradley (2004, 2012) and Conee (2006). According to the causal comparative account an event is harmful for someone if and only if its total causal consequence is worse for that person than its total prevention, where its total causal consequence is everything it causes to happen and its total prevention is what it causes not to happen. The causal comparative account solves some versions of the preemption problem, which my account does not, but the causal comparative account overcounts preventions. Suppose we wonder, what does an event cause *not* to happen? If the causal comparative account is to be distinguished from a straightforward counterfactual comparative account we *cannot* say that an event causes those events not to happen that otherwise would have happened. But then we seem to have to include events that would not have otherwise happened in the total prevention, and this results in too many preventions. By stealing the golf clubs, does the Joker cause Robin not to win the Masters? Robin would not have won anyway, but this can't matter on the causal comparative account (Bradley 2012: 409–410). On HAM, we do not need to know what an event caused not to happen to determine whether it is a harm. Instead, we need to ask (i) whether the person is worse off than he otherwise would have been had the event not occurred, and (ii) whether the event makes a person occupy his actual well-being level.

Finally, HAM does not limit failures to benefit (harm) to those events that *allow* a person to occupy their level of well-being in the world in which the events occur. This is because one might think that an event can be a failure to benefit (harm) when it (a) is *irrelevant* to the upshot (where relevance at least requires dependence) but (b) is such that had it not occurred S 's well-being would have been higher. This sort of case might very well be impossible, because, whatever relevance amounts to, if a person's level of well-being counterfactually depends on an event's occurrence, this entails that the event is *relevant* to the person's level of well-being. Still, I see no need to rule out this possibility in formulating HAM.

³⁸ It does not solve the preemption problem for the counterfactual comparative account of harm. But see Feit (2015) for an ingenious reply to the preemption problem. Perhaps some combination of the account offered here and Feit's account would both solve the preemption problem and successfully distinguish between harms and failures to benefit.

which the boulder is dislodged and (2) Victor is better off in the nearest possible world in which James does not dislodge the boulder. (2) is stipulated to be true. We can suppose that Victor occupies a lower well-being level for two years after he is maimed by the boulder than he would have occupied during those two years had the boulder not struck him. The important question, then, is whether dislodging the boulder *makes* Victor occupy this well-being level at the world at which the boulder is dislodged. To simplify matters, suppose that hedonism is true and so Victor's well-being level at a time is a function of the sum of the units of pleasure he enjoys at that time minus the sum of the units of pain he suffers at that time. Surely Victor's well-being level for the two-year period is one of the upshots of the boulder striking him. We can make use of Woollard's analysis to explain why it is that James's dislodging the boulder *makes* Victor occupy this level of well-being. First, the fact that James dislodged the boulder is a positive fact about James's behavior, so it is a substantial fact. Moreover, this fact about James's dislodging the boulder is relevant to Victor's occupying his level of well-being for two years through a complete chain of positive (and hence substantial) facts. Some of these facts are, say, that the boulder struck Victor in such-and-such a way, that Victor's bones were broken in the ways they were broken, and that Victor experienced severe pain and limited mobility for two years. Because it is stipulated that Victor is better off in the nearest possible world in which James does not dislodge the boulder, and because James *makes* Victor occupy his well-being level for the two years after the boulder strikes him, James's dislodging the boulder is a harm for Victor.

On the other hand, HAM counts Bob's driving away as a mere failure to benefit Victor. Bob's driving away is relevant to Victor's level of well-being through at least one negative fact (the absence of the car in the path of the boulder); the fact that the car is absent does not violate normal presuppositions; the resources required to sustain the presence of the car in the path of the boulder, which would have prevented Victor's occupying his level of well-being for two years after the boulder strikes him, do not belong to Victor or someone else who does not want him to occupy his level of well-being. Because it has these features, Bob's driving the car away is not relevant to Victor's occupying his level of well-being through a complete chain of substantial facts, and so it is not part of the sequence leading to Victor's occupying his level of well-being. Therefore, Bob's driving away does not make, but merely allows, Victor to occupy his level of well-being, and so it counts as a failure to benefit. HAM easily handles Drive Away 4. Let's return to the problem cases for CCA.

First, recall Robin's Clubs⁺: Batman decides to keep the clubs after swinging them a few times. Intuitively, Batman does not harm Robin by taking a few swings with the clubs. HAM can accommodate this intuition. Let us assume that Batman's swinging the clubs is relevant, in Woollard's thin sense of 'relevant', to Robin's level of well-being. According to HAM, Batman's swinging the clubs counts as a harm for Robin only if a fact about this action is relevant to Robin's level of well-being through a chain of substantial facts. Swinging the clubs is itself a substantial fact because it is a positive fact. However, Batman's swinging the clubs is relevant to Robin's well-being level through at least one negative fact: the fact that the clubs are absent, say, during his next round of golf. If this much is correct, then we need to

ask whether the negative fact that the clubs are absent meets the conditions under which a negative fact counts as a substantial fact. It must either violate normal presuppositions or it must be a negative fact about the absence of a barrier where the resources required to sustain the barrier belong to someone other than the agent (someone who does not want the upshot to occur). I suggest that the fact that the clubs are absent satisfies neither of these conditions for being a substantial fact in the case of Robin's Clubs⁺. The presence of a new set of gifted clubs is not like the presence of oxygen. It is not a background condition that we take for granted. Nor is fact that the clubs are absent a fact about the absence of a barrier, where the resources required to sustain the barrier belong to someone other than the agent who removed it. Setting aside for now the question about whether the clubs are a "barrier" to any upshot, the clubs are *Batman's*. Therefore, the fact about their absence is not substantial. Therefore, Batman's taking a few swings with the clubs is not part of the sequence leading to Robin's well-being level, and Batman's taking a few swings is not a harm for Robin.

HAM can handle preventative harm as well. In Joker's Removal, Batman has delivered golf clubs to Robin. The Joker removes them from Robin's porch just before Robin opens the door to retrieve them. It seems that the Joker (preventatively) harms Robin. Robin is better off in the nearest possible world in which the Joker does not remove the clubs. The important question is whether the Joker's removing the clubs counts as *making* Robin occupy his level of well-being in the world in which the Joker steals the clubs. This question is reducible to the question, Is the fact that the Joker removes the clubs a part of the sequence leading to Robin's level of well-being? That the Joker removes the clubs is a positive fact and it is relevant to Robin's level of well-being through a chain of facts at least one of which is negative: that the clubs are absent, say, during his next round of golf. Does this negative fact satisfy one of the two conditions under which a negative fact counts as substantial? I believe that it satisfies both conditions. First, the fact that the clubs are absent violates our normal presuppositions. To see this, note that it is not the same fact in Joker's Removal as the fact about their absence in Robin's Clubs⁺. The relevant negative fact in Joker's Removal is the fact that the clubs *that Batman gave to Robin* are absent. The absence of gifted clubs during Robin's next round of golf violates normal presuppositions. Normally, it is taken for granted that, once one is the recipient of a gift, that gift will remain in one's possession until another transfer occurs. Their absence is highly unexpected. Thus, the clubs' absence violates our normal presuppositions in Joker's Removal. Second, if we make what I take to be a friendly amendment to Woollard's analysis, the fact that the clubs are absent also satisfies Woollard's second condition under which a negative fact is substantial. The relationship between the absence of the clubs and Robin's level of well-being is strikingly similar to the relationship between the removal of a barrier and the upshot that results from its removal. I propose the following amendment to Woollard's analysis of the conditions under which a negative fact counts as a substantial fact.

Absences as Substantial A fact is *relatively substantial* if and only if it is a negative fact about the absence of a barrier to U or about the absence of a material condition that is necessary for

the provision of U, where sustaining the barrier or the material condition necessary for the provision of U requires the use of resources belonging to someone who did not remove the barrier or material condition who does (not) want U to obtain.³⁹

When we apply Absences as Substantial we see that, in Joker's Removal, the fact that the clubs are absent is a relatively substantial fact. Sustaining the material condition for the clubs to remain in Robin's possession requires the use only of resources belonging to Batman (or Robin, depending on when you think the clubs begin to belong to Robin). But Batman does not remove the clubs, nor does he want them removed. The Joker's removal of the clubs is thus relevant to Robin's well-being level through a relatively substantial negative fact about the absence of the clubs. Therefore, the Joker counts as *making* Robin occupy his well-being level in the world in which the Joker removes the clubs. Just as the removal of Victor's car from the path of the boulder counts as making the boulder crush him, because it removes a barrier belonging to Victor that would have prevented him from being crushed, on my analysis, the removal of the clubs counts as making Robin occupy his actual well-being level in virtue of removing a necessary material condition for the welfare level that Robin would otherwise have enjoyed. Therefore, HAM, in conjunction with my small amendment to Woollard's analysis, entails that the removal is a harm for Robin.

To sum up: the difference between mere failures to benefit and preventative harms is that preventative harms, but not failures to benefit, make a person occupy their actual well-being level. A preventative harm counts as making a victim occupy his actual level of well-being because it is relevant to a victim's actual well-being level through a negative relatively substantial fact about the absence of a material condition necessary for that victim to occupy a higher well-being level than his actual one. The negative fact about the absence of a material condition is relatively substantial when either (a) sustaining that material condition requires the use of resources belonging to someone who did not remove the material condition, or (b) the fact violates normal presuppositions.

One might object that addressing HAM's problem cases by appealing to Woollard's analysis means giving up the ontological neutrality of our account of harm. A successful account of harm will provide conditions under which *any* event is a harm, but HAM requires identifying whether the material condition for a benefit belongs to an agent when determining whether an event counts as a harm. Does this mean that HAM is silent about harms that are not actions? I do not think so.⁴⁰ We can ask, for *any* removal of a material condition for some welfare benefit, whether (a) the event is an action performed by the agent to whom the material condition belongs who wants the upshot to obtain, and whether (b) the removal of the condition violates normal presuppositions. Suppose that life-saving medical

³⁹ Here I have removed Woollard's second sufficient condition regarding non-need based claims, because this condition is not important for my purposes.

⁴⁰ Woollard's original analysis violates ontological neutrality, but my friendly amendment does not.

supplies are airdropped to stranded hurricane victims. Later, a flash flood washes them away. The fact that the supplies are absent is a negative fact about the absence of a material condition necessary for the hurricane victims to occupy a higher well-being level than their actual one, and sustaining this material condition requires the use of resources (the medical supplies) belonging to someone (either the hurricane victims or the aid organization) who did not remove the material condition. Therefore, HAM counts it as a harm. There is no threat to ontological neutrality here.

It is worth noting two further advantages of HAM over CCA. First, HAM distinguishes harms from failures to benefit that involve *failing to prevent harm*. Steve will be struck by an oncoming bus unless I push him out of the way. I choose instead to call for help. Had I not called for help I would have pushed Steve out of the way. Steve is better off in the nearest possible world in which I do not call for help. Because Steve is better off in the nearest possible world in which I do not call for help, CCA entails that my calling for help is a harm for him. Intuitively, calling for help is not a harm. HAM respects this judgment. My calling for help does not *make* Steve occupy his level of well-being, because it is not part of the sequence leading to his level of well-being. The fact that I called for help is relevant to Steve's well-being level through a *non-substantial* negative fact—the fact that I did not push Steve out of the way. This negative fact is not substantial because: (a) it is not a fact about the absence of a barrier or a fact about the absence of a material condition necessary for Steve to occupy a higher well-being level than his actual well-being level, where sustaining the barrier (material condition) requires the use of resources belonging to someone who did not remove the barrier (material condition); and (b) the fact that I did not push Steve out of the way does not violate normal presuppositions. Therefore, though Steve is worse off than he would have been had I not called for help, calling for help merely allows Steve occupy his level of well-being.

Second, HAM avoids the implication that events can be harmful merely because of what they signify. If an x-ray reveals that a patient has a torn ACL, it is reasonable to say that it is a bad thing that the x-ray revealed torn ligaments.⁴¹ It is not reasonable to say that the x-ray revealing torn ligaments is a *harm* for the patient. This is because the x-ray is bad only in virtue of what it signifies. Yet CCA entails that the x-ray revealing torn ligaments is a harm for the patient, because the patient is better off in the nearest possible world in which the x-ray does not reveal torn ligaments.⁴² HAM does not have this odd implication. Because the x-ray does not *make* the patient's level of well-being lower than it otherwise would have been, it is not a harm.

⁴¹ This example is Bradley's (1998: 118).

⁴² This problem cannot be avoided simply saying that backtracking counterfactual claims are all false. Backtracking counterfactuals are claims about how things would be earlier if later things were different. If backtracking counterfactual claims are all false, we can imagine that, rather than revealing information about past events, x-rays worked by reliably predicting future states of affairs on the basis of the best presently available evidence. This would be a remarkable machine indeed, but I am not convinced that it is metaphysically impossible.

To sum up, the central challenge addressed in this section is that CCA counterintuitively entails that Batman harms Robin by taking a few swings with the clubs and that Archie harms Jughead by unsticking the dollar, that I harm Steve by calling for help, and that the x-ray is harm for the patient. HAM avoids these counterintuitive implications, and it does so in a principled way. It thereby promises greater extensional adequacy than CCA, and it preserves the intuitive moral asymmetry between harms and failures to benefit. For those who take seriously the distinction between harms and failures to benefit, HAM is an improvement.

4 Objections

In this section I consider four objections that might be pressed against HAM.

4.1 Killing

HAM says that harms make a victim occupy his level of well-being. If HAM is to count an event as a harm, then the event's victim must possess a level of well-being, but a person does not exist after his death. Because non-existent individuals do not have levels of well-being, a dead person has *no* level of well-being at all. If a dead person does not possess a level of well-being then an event cannot make a dead person occupy a level of well-being. If all of this is right, then HAM entails that a killing is not a harm for its victim because, even if the victim would have otherwise been better off, the killing does not make the victim occupy a level of well-being.

One reply to this objection would be to insist that the dead possess levels of well-being. In general, individuals can have properties at times when they do not exist. For example, the dead can be eulogized.⁴³ Perhaps a dead person can also possess a level of well-being. If they do possess a level of well-being, this must be in virtue of the fact that certain states of affairs fail to obtain. I'm not sure this is a problem. A person in a coma arguably has a level of well-being of zero, and this is at least partly true in virtue of the fact that he is, say, *neither* enjoying pleasure nor suffering pain. Thus, there is no special problem with saying either that an individual can have a level of well-being at times when he does not exist, and that states of affairs that fail to obtain at those times can ground that level of well-being.⁴⁴ If a dead person possesses a well-being level of zero, then perhaps we can use Woollards's account to explain how it is that a killing event makes a person occupy this level of well-being. For example, suppose that I am shot in the head and killed. Suppose also that

⁴³ Bradley (2009) uses this example in defending the claim that the dead can possess a level of well-being.

⁴⁴ Bradley (2009) gives a similar argument. See also Feit (2016) who appeals to Kit Fine's distinction between propositions being true *in* a possible world and propositions being true *at* a possible world. Feit argues that while it is false that I have a (zero) level of well-being *in* some possible world at a time at which I do not exist, it might be true that I have a (zero) level of well-being *at* some possible world at a time at which I do not exist, where my having a level of well-being of zero is grounded in facts about what I *lack* at that time at that world (e.g., any experiences of pleasure or pain). But see Carlson and Johansson (2018) for a reply to Feit.

the fact that certain pleasant conscious experiences are absent at a time after my death is one of the grounds of my having a zero-level of well-being at that time. Surely, then, the fact that I was shot in the head is relevant to my having a zero-level of well-being *through* this negative fact about the absence of pleasant conscious experiences. But this negative fact is itself grounded in a fact about the absence of certain brain states. We are now in a position to apply Woollard's amended analysis of the conditions under which negative facts are substantial facts in order to determine whether the shooting event counts as *making* me occupy a well-being level of zero: I propose that the negative fact about the absence of certain brain states at some time after my death is a negative fact about the absence of a material condition necessary for the provision of a well-being level at that time that is higher than my actual well-being level. Had certain brain states obtained at that time, my level of well-being would have been higher at that time. I suggest further that sustaining this material condition requires the use of resources (my brain) belonging to someone (me) who did not remove the material condition. The *shooter* removed a material condition belonging to *me*, the victim of the shooting. Because my brain is my resource, the absence of certain brain states of mine is a relatively substantial fact through which the event of shooting me is relevant to my well-being level of zero. Shooting *me* *makes* me possess a zero level of well-being.

But suppose that the dead do not have levels of well-being. What can a defender of HAM say about the harm of killing? Perhaps we can say that killing a person eliminates him, the subject of his life, capping his total amount of well-being at whatever it was when the person died. How can the elimination of a person count as *making* that person's total lifetime well-being be what it is in the world where the event occurs? Again, I want to suggest that the fact that I was shot in the head is relevant to the upshot *my having amount n of lifetime well-being* through a negative fact about the absence of pleasant conscious experiences. But this negative fact is itself grounded in a fact about the absence of certain brain states. From here we can again apply Woollard's analysis of the conditions under which a negative fact is a substantial fact. When we do this, we will again find that the absence of certain brain states is a substantial fact. Therefore, if the dead do not possess levels of well-being, HAM can account for the harm of death if HAM is recast in terms of 'total amount of lifetime well-being' rather than 'levels of well-being'. I continue to use the term 'level of well-being' but assume this language could be replaced by 'total amount of lifetime well-being'.⁴⁵

⁴⁵ There is a puzzle here about suicide. A momentarily depressed person who kills himself seems to harm himself, but my explanation of the harmfulness of killing seems unable to explain why. After all, momentarily depressed suicide victim removes material conditions for the provision of his well-being belonging only to *him*. Answering this worry requires providing an account of the conditions under which a person's body counts as a resource belonging to *him*. I can only briefly, controversially, and perhaps incoherently suggest that an irrationally suicidal person's body (and I assume that the momentarily depressed person is one such example) does not belong to them at the time of their illness. This is precisely why certain coercive interventions are not considered violations of bodily autonomy when they are designed to prevent a person with severe mental illness from harming themselves.

4.2 Harming by omission

Another concern is that HAM cannot account for harming by omission. Consider parents who neglect to feed their infant, or the lifeguard who sits idly by while a weak swimmer drowns. These events seem to be harms, but they are also mere allowings. These cases are difficult to adjudicate because they involve a prior moral commitment on the part of the agent to protect some individual, and this commitment is violated by the purportedly harmful omission. Parents take on a special moral obligation by taking an infant into their custody. Lifeguards take on a special moral obligation to beachgoers when they take the job. Because they involve violations of special moral obligations, failures to benefit are morally equivalent to (wrongful) harms in these cases. The beachgoer has a special claim on the lifeguard such that the lifeguard wrongs the beachgoer by failing to benefit him (by failing to save him). Finally, these omissions all involve absences that violate normal presuppositions. The fact that a newborn has not been fed by his parental guardians, and the fact that an on-duty lifeguard does not make an effort to save a drowning swimmer are facts that violate our normal presuppositions. The negative fact that the newborn has not been fed by his parental guardian is therefore a substantial fact relevant to the starvation of their child. The parents' behavior (e.g., removing him from the care of the hospital staff and placing him in a crib by himself) is relevant to the starvation of their child through this substantial fact. So, the parents' behavior is a good candidate to be a part of a sequence leading to the starvation of their infant. Thus, though HAM entails that the neglect itself is not a harm, the neglect is the sort of negative fact capable of rendering harmful the parents' earlier actions.

One might further press the issue of harming by omission. David Boonin gives an example of a harmful event in which: (i) the event merely allows a subject's well-being to be lower than it would have been had the event not occurred, and (ii) the agent has no special obligation to the subject. "If I decline to give you a pill that I have in my pocket that would prevent you from suffering a great deal of pain, ... I think most people would find it perfectly plausible and even natural to say that I have harmed you by withholding the medicine."⁴⁶ It is doubtful that cases of this sort can decide the issue, because they too closely resemble the cases whose status is up for debate. The central issue under contention is whether, for instance, a pharmaceutical company that 'withholds' medication from people who are unable to pay for it *harms* those people or merely fails to benefit them; whether withholding medication for a fatal disease is metaphysically (and morally) the same as making a person contract it in the first place.

4.3 Bestowing 'pure' benefits v. preventing harm

Seana Shiffrin has argued that appealing to the making/allowing distinction will not enable the counterfactual comparative account of harm to make sense of all of the asymmetries between harms and benefits that must be accommodated by a plausible

⁴⁶ Boonin (2014: 53, n2). Feit (2017) discusses the example as well.

theory.⁴⁷ Suppose that sufficient compensation for a broken leg is \$150,000. That is, if someone suffers a broken leg, being paid \$150,000 leaves one exactly as well off as one would have been had he not suffered the injury. Shiffrin's view is that you have a stronger reason to prevent a broken leg than to pay someone \$150,000 because preventing the injury benefits someone by preventing *harm* whereas paying someone \$150,000 merely provides a *pure* benefit, where 'pure benefit' refers to any benefit that is not also a prevention of harm. But HAM cannot explain this asymmetry concerning an agent's moral reasons. According to HAM, neither refraining from preventing the break nor refraining from giving someone \$150,000 counts as a harm.

To understand Shiffrin's objection, one must understand her starting point for a theory of harming. She claims that a theory of harming "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."⁴⁸ On her view, these states are harmful regardless of how they compare to alternatives and regardless of how they came about. Having a broken leg is a harmful condition to be in but (merely) lacking an additional \$150,000 is not. HAM is not an account of the harmfulness of conditions. It cannot explain why failing to prevent a broken leg is morally worse than failing to give someone an additional \$150,000, because both involve allowing the 'victim' to be worse off to the same extent than they otherwise would have been. However, HAM need not explain this asymmetry. Whether an action is wrong depends factors other than the action's status as a harm or mere failure to benefit. For example, one moral difference between preventing a broken leg and providing \$150,000 dollars is that people generally have negative *rights* not to have their legs broken that are stronger than any positive right to receive a monetary award. The varying strength of rights can explain the comparative strength of our reasons to prevent a broken leg and to provide \$150,000.

Shiffrin also argues that it is generally worse to allow harm than to cause the removal of a comparably sized pure benefit. For example, it is worse to allow someone's leg to be broken than to prevent someone from receiving full funding for her research project. HAM entails that allowing the break is a failure to benefit, whereas interrupting the funding is a harm, so interrupting is *prima facie* morally worse. But HAM does not entail the further claim that interrupting the funding is morally worse *overall* than failing to prevent the broken leg; for HAM is not an account of the overall comparative moral importance of various sorts of harming and benefitting events. In fact, it does not say anything at all about the comparative moral importance of harming and benefitting. It is an account of which events count as harms and failures to benefit.⁴⁹

⁴⁷ Shiffrin (2012: 375).

⁴⁸ Shiffrin (2012: 361).

⁴⁹ Hanna (2015: 8) and (Bradley 2012: 410) have argued that a theory of harm need not account for all of the moral differences between types of harms and benefits. I agree.

4.4 Preventing preventions

Matthew Hanser has argued that the counterfactual comparative account cannot distinguish between ‘ordinary’ harms and harms that prevent someone from being prevented from suffering a harm. Hanser’s example involves the loss of one’s sight (2008: 427–429). Suppose that Nefarious Villain 1 throws acid into my eyes, which makes me lose my sight. A doctor is on his way to perform a surgery that prevents me from losing my sight, but Nefarious Villain 2 slashes the doctor’s tires, preventing the doctor from performing the surgery. The problem for CCA, Hanser claims, is that

From the comparative point of view...there is no difference between the ordinary harm of losing one’s sight and the preventative harm of being prevented from being prevented from losing one’s sight. The two have identical comparative profiles. So while the counterfactual comparison analysis can distinguish some preventative harms from ordinary harms, it cannot draw the distinction in all cases (2008: 429).

I take it that Hanser’s complaint is twofold: first, that the counterfactual comparative account cannot draw a metaphysical distinction between preventing preventions and ‘ordinary’ harms and, thus, second, that it cannot draw a moral distinction between them. A defender of HAM can answer the objection in two ways. First, the purpose of an account of harm is to distinguish between harms and non-harms. It is unclear why an account of harm must also tell us which events are preventions, which are preventions of preventions, and which categories of harming are morally worse than others. Hanser’s objection may thus demand too much. Second, the foregoing discussion of the different ways that an event can count as part of a sequence leading to an upshot suggests a reply to Hanser’s objection. Throwing acid into my eyes and slashing the surgeon’s tires differ with respect to their role in the sequence leading to the loss of my sight: The fact that Nefarious Villain 1 threw the acid is relevant to the loss of my sight through a chain of substantial positive facts, whereas the fact that Nefarious Villain 2 slashed the surgeon’s tires is relevant to the loss of my sight through a (substantial) negative fact about the absence of a barrier to the loss of my sight. In addition to the distinction between making and allowing, there may be further morally relevant distinctions between *ways* of making upshots occur, where their classification depends on whether a fact about the event is relevant through a chain of purely positive facts or at least one negative fact. The moral difference between Nefarious Villain 1’s action and Nefarious Villain 2’s action may thus be explainable by positing a moral distinction between positive and negative facts. This reply is available to both HAM and CCA.

5 Concluding remarks

Though it has not been taken up in the literature, it should come as little surprise that ordinary intuitions about harm track the distinction between making upshots happen and allowing upshots to happen. The moral significance of the distinction between harming and failing to benefit is critical to distinguishing consequentialist and deontological moral theories. Non-consequentialists, many of whom share a commitment to the existence of moral constraints on the promotion of the good, can accommodate the intrinsic moral significance of the fact that an action is harmful. Indeed, the notion of harm figures centrally in the content of moral constraints. To take one example, consider W.D. Ross's discussion of the *prima facie* duties of beneficence and non-maleficence.

But even when we have come to recognize the duty of beneficence, it appears to me that the duty of non-maleficence is recognized as a distinct one, and as *prima facie* more binding. We should not in general consider it justifiable to kill one person in order to keep another alive, or to steal from one in order to give alms to another (Ross 1930: 22).

If we were to count all failures to benefit as harms, and we were to accept common prohibitions against harming, then the moral demands on agents would be more burdensome than ordinary morality allows. Standard prohibitions against harming would recommend a great deal of interference if they applied to failing to benefit. Take as an example John Stuart Mill's harm principle.

The only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. (Mill 1978: 9)

The harm principle is often invoked as a justification for state interference with certain forms of behavior. If we counted an action as a harm in any case where the agent would otherwise have benefitted a person, Mill's harm principle would justify coercive intervention in too many cases. It might, for instance, recommend the use of coercive power to ensure that agents provide certain goods to individuals to whom they have no relationship. Mill surely meant his principle to apply more narrowly than this. It is thus important for preserving non-consequentialist moral commitments to determine whether the distinction between harms and failures to benefit can be sustained, and in which cases. It makes a great deal of sense that this would ultimately require appealing to the making/allowing distinction, which has also proven essential to preserving non-consequentialist moral commitments (Woollard 2017).

In this paper I showed that non-counterfactual comparative of analyses of harm rely on the making/allowing distinction in distinguishing harming from failing to benefit. I then drew on this feature of these accounts to advance an account of harm that distinguishes harming from failing to benefit by restricting harms to those events that make a person—rather than merely allow him to—occupy his level of well-being at the world at which the event occurs, *and* are such that the person is

worse off than he would have been had they not occurred. This is evidence that an appeal to the distinction between making and allowing is essential to formulating an account of harm that respects the distinction between harms and failures to benefit. Of course, this means that if the former distinction is untenable, then so is latter. This is further reason to settle the debate about the moral significance of the distinction between making and allowing.

Acknowledgements I owe a significant debt to Neil Feit, Molly Gardner, Jens Johansson, Stephen Kershnar, Justin Kloksiem, and Michael Tooley for helpful written comments on earlier drafts of this paper. Thanks to Matthew Hanser for a particularly illuminating conversation about these issues. I am also grateful to participants at the 2016 Bled Ethics Conference for their valuable feedback. Finally, this paper was significantly improved by a challenging set of comments from an anonymous referee for *Philosophical Studies*.

References

- Bennett, J. (1995). *The act itself*. Oxford: Oxford University Press.
- Boonin, D. (2014). *The non-identity problem and the ethics of future people*. Oxford: Oxford University Press.
- Bradley, B. (1998). Extrinsic value. *Philosophical Studies*, 91(2), 109–126.
- Bradley, B. (2004). When is death bad for the one who dies? *Noûs*, 38(1), 1–28.
- Bradley, B. (2009). *Well-being and death*. Oxford: Oxford University Press.
- Bradley, B. (2012). Doing away with harm. *Philosophy and Phenomenological Research*, 85(2), 390–412.
- Carlson, E., & Johansson, J. (2018). Well-being without being?: A reply to Feit. *Utilitas*, 30(2), 198–208.
- Conee, E. (2006). Dispositions toward counterfactuals in ethics. In K. McDaniel, et al. (Eds.), *The good, the right, life and death* (pp. 173–187). Farnham: Ashgate Publishing.
- Feinberg, J. (1986). Wrongful life and the counterfactual element in harming. *Social Philosophy and Policy*, 4(1), 145–178.
- Feit, N. (2015). Plural harm. *Philosophy and Phenomenological Research*, 90(2), 361–388.
- Feit, N. (2016). Comparative harm, creation, and death. *Utilitas*, 28(2), 136–163.
- Feit, N. (2017). Harming by failing to benefit. *Ethical theory and moral practice*. <https://doi.org/10.1007/s10677-017-9838-6>.
- Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, 5, 5–15.
- Foot, P. (1977). Euthanasia. *Philosophy & Public Affairs*, 6(2), 85–112.
- Hanna, N. (2015). Harm: Omission, preemption, freedom. *Philosophy and Phenomenological Research*, 91(2), 1–23.
- Hanser, M. (2008). The metaphysics of harm. *Philosophy and Phenomenological Research*, 77(2), 421–450.
- Harman, E. (2004). Can we harm and benefit in creating? *Philosophical Perspectives*, 18(1), 89–113.
- Kloksiem, J. (2012). A defense of the counterfactual comparative analysis of harm. *American Philosophical Quarterly*, 49(4), 285–300.
- Lewis, D. (1973). Causation. *Journal of Philosophy*, 70, 556–567.
- Mackie, J. L. (1965). Causes and conditions. *American Philosophical Quarterly*, 2(5), 245–264.
- McMahan, J. (1993). Killing, letting die, and withdrawing aid. *Ethics*, 103, 250–279.
- Mill, J. S. (1978). *On liberty*. Indianapolis: Hackett Publishing Company.
- Perry, S. (2003). Harm, history, and counterfactuals. *San Diego Law Review*, 40, 1283–1313.
- Purves, D. (2015). The harms of death. In M. Cholbi (Ed.), *Immortality and the philosophy of death*. Lanham: Rowman and Littlefield.
- Purves, D. (2016). Accounting for the harm of death. *Pacific Philosophical Quarterly*, 97(1), 89–112.
- Rachels, J. (1975). Active and passive euthanasia. *New England Journal of Medicine*, 292, 78–80.
- Roberts, M. (1998). *Child versus childmaker: Future persons and present duties in ethics and the law*. Lanham, MD: Rowman and Littlefield.
- Ross, W. D. (1930). *The right and the good*. Oxford: Clarendon Press.

-
- Shiffrin, S. V. (1999). Wrongful life, procreative responsibility, and the significance of harm. *Legal Theory*, 5(2), 117–148.
- Shiffrin, S. V. (2012). Harm and its moral significance. *Legal Theory*, 18(3), 357–398.
- Stalnaker, R. C. (1968). A theory of conditionals. *American Philosophical Quarterly*, 98–112.
- Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist*, 59(2), 204–217.
- Woollard, F. (2008). Doing and allowing, threats and sequences. *Pacific Philosophical Quarterly*, 89(2), 261–277.
- Woollard, F. (2015). *Doing and allowing harm*. Oxford: Oxford University Press.
- Woollard, F. (2017). Double effect, doing and allowing, and the relaxed nonconsequentialist. *Philosophical Explorations*, 20(sup2), 142–158.