

Well-Being Counterfactualist Accounts of Harm and Benefit

Erik Carlson, Jens Johansson & Olle Risberg

To cite this article: Erik Carlson, Jens Johansson & Olle Risberg (2021) Well-Being Counterfactualist Accounts of Harm and Benefit, Australasian Journal of Philosophy, 99:1, 164-174, DOI: [10.1080/00048402.2019.1705363](https://doi.org/10.1080/00048402.2019.1705363)

To link to this article: <https://doi.org/10.1080/00048402.2019.1705363>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 02 Feb 2020.



Submit your article to this journal [↗](#)



Article views: 1792



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 2 View citing articles [↗](#)



Well-Being Counterfactualist Accounts of Harm and Benefit

Erik Carlson, Jens Johansson, and Olle Risberg

Uppsala University

ABSTRACT

Suppose that, for every possible event and person who would exist whether or not the event were to occur, there is a well-being level that the person would occupy if the event were to occur, and a well-being level that the person would occupy if the event were not to occur. Do facts about such connections between events and well-being levels always suffice to determine whether an event would harm or benefit a person? Many seemingly attractive accounts of harm and benefit entail an affirmative answer to this question, including the widely held Counterfactual Comparative Account (CCA). In this paper, however, we argue that all such accounts will be unsuccessful.

ARTICLE HISTORY Received 7 May 2019; Revised 5 December 2019

KEYWORDS harm; benefit; well-being; counterfactual comparative account of harm

1. Introduction

Suppose that, for every possible event e and person S who would exist whether or not e were to occur, there is a well-being level W_{S_e} such that S would occupy W_{S_e} if e were to occur, and a well-being level $W_{S_{\neg e}}$ such that S would occupy $W_{S_{\neg e}}$ if e were not to occur. Are facts about such connections between events and well-being levels always sufficient to determine whether a particular event would harm or benefit a particular person? Let us call an account of harm and benefit *well-being counterfactualist* if and only if it answers this question affirmatively. Many seemingly attractive accounts of harm and benefit are well-being counterfactualist in this sense. However, proceeding by first discussing a very simple well-being counterfactualist account, and then moving on to progressively more complex ones, we will arrive at the conclusion that no such account is acceptable.

According to the most widely held and discussed account of harm and benefit, often called the *counterfactual comparative account* (CCA), an event harms (benefits) a person just in case her well-being level would have been higher (lower) if the event had not occurred.¹ Since CCA is well-being counterfactualist, it is among the targets of our criticism. While there are some prominent objections to CCA in the literature, it is controversial how compelling these are. Our rejection of CCA, however, is

¹ Although we shall follow established usage and use 'CCA' to refer to this particular account, it should be noted that many rival accounts are no less counterfactual and comparative.

mainly based on considerations that have not previously been fully articulated. If our general conclusion is correct, certain initially attractive ways of modifying CCA are doomed to fail as well.

2. Preliminary Notes

We are interested in accounts of *overall*, rather than *pro tanto*, harm and benefit. While an event might be both *pro tanto* harmful and *pro tanto* beneficial to a person, and might be both overall harmful (beneficial) and *pro tanto* beneficial (harmful) to a person, we shall assume that no event is both overall harmful and overall beneficial to one and the same person.

Sometimes a distinction is made between *intrinsic* and *extrinsic* harm (see, for instance, Bradley [2012: 392] and Klocksiem [2012: 289]). An event is intrinsically harmful to the extent that it is harmful because of its intrinsic properties. An event is extrinsically harmful to the extent that it is harmful because of its extrinsic properties—in particular, those that concern its relations to other events, such as its effects. A corresponding distinction can be made regarding benefit. We shall understand overall harm and benefit as including both intrinsic and extrinsic harm and benefit. Thus, an event might be intrinsically harmful (beneficial) but overall beneficial (harmful), due to its extrinsic properties. Similarly, an event might be extrinsically harmful (beneficial) but overall beneficial (harmful), because of its intrinsic properties.

According to most advocates of well-being counterfactualist accounts, what determine overall harm and benefit are the *lifetime* well-being levels to which the relevant events would give rise. An alternative view appeals instead to the well-being level that the person would have *after* each relevant event (see Johansson [forthcoming]). Yet further alternatives are possible. While we shall adopt the ‘lifetime’ conception in our formulations of specific well-being counterfactualist accounts, nothing in our discussion hinges on this.

We shall make no particular suppositions about the components of well-being (for example, whether they are episodes of pleasure, or desire satisfactions, etc.). As is customary in the debate, however, we presuppose that well-being is intrapersonally measurable on at least an interval scale.

3. Absolutism

We will make the probably unrealistic, but simplifying, assumption that every possible event counterfactually determines a lifetime well-being level for every person who would exist if the event were to occur. Thus:

Assumption 1. For any possible event e and person S who would exist if e were to occur, there is a lifetime well-being level W_{Se} such that S would occupy W_{Se} if e were to occur.²

This assumption follows from, but does not entail, the thesis of ‘counterfactual determinism’, according to which, for every possible event e , there is a unique possible world w such that if e were to occur then w would obtain.

² Throughout the paper, we assume that it is legitimate to quantify over merely possible events and persons. While this is not needed for our purposes, it simplifies the exposition. As regards well-being levels, we assume that these are abstract entities that exist in every possible world.

A very simple well-being counterfactualist theory of harm and benefit is the following:

Absolutism. There is a lifetime well-being level W such that, for any possible event e and person S , e would harm S iff $W > W_{S_e}$, and e would benefit S iff $W_{S_e} > W$.³

Absolutism can be quickly dismissed. For any actually occurring event e and any actual person S , W_{S_e} is identical to S 's actual lifetime well-being level. Hence, Absolutism implies that every person with an actual lifetime well-being level below W is harmed by every actual event, and that every person with an actual lifetime well-being level above W is benefited by every actual event. This implication is unacceptable, for two interrelated reasons. First, even a life that is very good for the person living it may surely contain some events that harm her; similarly, even a life that is very bad for the person living it may surely contain some events that benefit her. Second, let us call an event that neither harms nor benefits a person S *neutral* for S . It is clear that, for any actual person, a great many actual events are neutral for her. Consider, for instance, a microphysical event occurring today in a star millions of light-years from Earth. It is evident, we take it, that such an event is neutral for every currently living human person.

Absolutism thus violates the following adequacy condition for a theory of harm and benefit:

Adequacy Condition 1. For any lifetime well-being level W , a person S 's occupying W is compatible with there being actual events of each of these three categories—harmful for S , beneficial for S , and neutral for S .

Adequacy Condition 1 implies that Absolutism states neither necessary nor sufficient conditions for harm and benefit.

4. The Counterfactual Comparative Account

The failure of Absolutism shows that there is no universal 'baseline' well-being level that, for any possible event e and person S , determines whether e harms or benefits S . The antithesis to Absolutism is extreme relativism concerning baselines, positing an individual baseline for each e and S . A natural idea is to let the baseline be the lifetime well-being level that S would occupy if e were not to occur. Let us thus make the following assumption:

Assumption 2. For any possible event e and person S who would exist if e were not to occur, there is a lifetime well-being level W_{S_e} such that S would occupy W_{S_e} if e were not to occur.

A clarificatory comment is needed: W_{S_e} is assumed to be S 's lifetime well-being level in the non- e -world that is nearest as seen from *the nearest e-world*, rather than as seen from the *actual* world (cf. Carlson [2019: 796]). These worlds might be different, if the actual world is a non- e -world.

We can now state the following account:

³ Note that Absolutism is distinct from most 'non-comparative' theories of harm discussed in the literature. For example, Absolutism neither follows from nor entails the simple causal theory that an event e harms an agent S iff e causes something that is intrinsically bad for S . This is because the claim that e causes something intrinsically bad for S does not entail that S occupies any particular lifetime well-being level; and not even together with the assumption that W is 0 or lower does the claim that $W > W_{S_e}$ entail that e causes something intrinsically bad for S .

The Counterfactual Comparative Account (CCA). For any possible event e and person S , e would harm S iff $W_{S-e} > W_{Se}$, and e would benefit S iff $W_{Se} > W_{S-e}$.

CCA is the most widely held and discussed theory of harm and benefit.⁴ The theory is simple, elegant, and fulfils Adequacy Condition 1, since e comes out as neutral if $W_{Se} = W_{S-e}$. But CCA also faces notorious difficulties. One is known as ‘the preemption problem’.⁵ Suppose that Jill hits Jack with her fist, breaking his nose. Had Jill not hit Jack with her fist, she would instead have hit him with a baseball bat. This would have injured Jack even more, and would have resulted in a lower lifetime well-being level for him. In this case of what appears to be ‘preemptive harm’, CCA implies that Jill’s hitting Jack with her fist benefits him. Many would find this conclusion plainly false; on the contrary, they would claim, Jill’s action harms Jack. There are also cases of what is seemingly ‘preemptive benefit’. Suppose that Mary invites Mike to a party, which he enjoys very much. Had Mary not invited Mike to the party, however, she would instead have invited him to a concert, which he would have enjoyed even more. His lifetime well-being level would, as a result, have been higher. CCA implies that Mary’s inviting Mike to the party harms him. Many would find it obvious that it does not; rather, they would maintain, it benefits him.

However, it is controversial whether the preemption problem does in fact establish the falsity of CCA [Klocksiam 2012; Boonin 2014: 62–3; Feit 2015]. We are inclined to think that it does. But, instead of relying on this claim, we shall present another argument against CCA. To begin with, we shall assume that, for every possible choice situation, there is a set of possible actions that constitute the options for the agent in that situation:

Assumption 3. For any possible choice situation C_S , there is a finite set O_{CS} of mutually exclusive and jointly exhaustive possible actions that are performable by the agent S in C_S . Call such a set an *option-set*.

The performable actions are those that the agent *can* perform, in the sense at issue in the principle that ‘ought implies can.’ It is commonly assumed—for example, in standard formulations of consequentialism and other ethical theories—that there is a set of such actions. While there might be some situations in which the agent could not have done anything other than what she did (and so situations in which O_{CS} has only one member), we take it that O_{CS} typically has at least two members.

Using the concept of an option-set, we can formulate the following very plausible adequacy condition for theories of harm and benefit:

Adequacy Condition 2. Let possible actions a and a^* be members of the same option-set O_{CS} . If W_{Sa} is much higher than W_{Sa^*} and there are no other relevant differences between a and a^* , then a would harm S only if a^* would harm S , and a^* would benefit S only if a would benefit S .

Adequacy Condition 2 does not require for its plausibility the ‘*much higher*’ (rather than merely ‘higher’) clause. It is still worth including this clause, as it makes Adequacy Condition 2, as well as our arguments below, even more difficult to resist.

That there are ‘no other relevant differences’ between a and a^* is here stipulated to mean that a and a^* do not differ with regard to factors such as causation,

⁴ For defences of CCA, see, for example, Parfit [1984: 69], Bradley [2009], Klocksiam [2012], Boonin [2014], Feit [2015, 2016, 2019], and Timmerman [2019].

⁵ See, for example, Norcross [2005], Bradley [2012], Gardner [2015], Hanna [2016], and Johansson and Risberg [2019].

rights-violations, the agent's motives or intentions, or other factors that opponents of well-being counterfactualist accounts have sometimes held to be relevant to harm and benefit (see, for example, Woodward [1986], Harman [2009], Gardner [2015], and Purves [2019]). Without this *ceteris paribus* clause, Adequacy Condition 2 might be rejected (although not by well-being counterfactualists) on, for instance, the ground that, even if W_{Sa} is much higher than W_{Sa^*} and a^* would not harm S, it can still be true that a would harm S so long as a but not a^* would *cause* S to occupy the relevant well-being level. With the *ceteris paribus* clause in place, however, Adequacy Condition 2 is eminently plausible. What it states, less formally, is simply that, in a given situation, a harmful action cannot leave the person much better off than a harmless action, and a beneficial action cannot leave the person much worse off than a non-beneficial action—assuming, again, that there are no other relevant differences (to do with causation, rights-violations, etc.) between the respective actions. This strikes us as a platitude that also helps to explain the idea that harm is prudentially relevant. It is hard to see how we could have prudential reasons to avoid harmful actions and to choose beneficial ones if this principle were false (since the harmful actions could in that case leave us at much higher well-being levels than the beneficial ones, even if there are no other relevant differences between them).

The problem for CCA is that it is incompatible with Adequacy Condition 2. Suppose that Beth has the following four options—to visit a fantastic resort; to visit a moderately nice resort; to visit a terrible resort; or to voluntarily go to prison, where she would be confined for life. The resulting lifetime well-being levels for Beth are ordered accordingly, with very large differences between them. As seen from the nearest possible world where Beth goes to the moderately nice resort, the nearest world where she does not do so is one where she goes to the extremely nice resort. As seen from the nearest world where she goes to the terrible resort, the nearest world where she does not do so is one where she goes to prison. (While going to the terrible resort is an option for Beth, she would not perform that action unless she had lost her preference for pleasurable experiences, in which case she might just as well go to prison.)

Under these assumptions, CCA implies that Beth's going to the moderately nice resort would harm her, while going to the terrible resort would benefit her. This result is problematic. For since Beth's lifetime well-being level would be much higher if she went to the moderately nice resort than if she went to the terrible resort (and there are no other relevant differences between these actions), Adequacy Condition 2 requires that the former action would harm her only if the latter action would harm her. This is incompatible with CCA.

It is important to note that our objection is not that CCA faces both the 'preemptive harm' problem and the 'preemptive benefit' problem in one and the same case.⁶ A 'preemptive harm' objection to CCA, based on the present case, would claim that going to the moderately nice resort would not harm Beth (contrary to CCA); a 'preemptive benefit' objection would claim that going to the terrible resort would not benefit her (contrary to CCA). Our objection assumes neither of these claims. Still less, of course, does it assume their conjunction. Our objection is, instead, that, since going to the moderately nice resort would leave Beth much better off than going to the terrible resort (and these actions belong to the same option-set, and there are no other relevant differences between them), it cannot *both* be that going to the moderately nice resort

⁶ Thanks to an anonymous reviewer for prompting us to clarify this.

would harm Beth *and* that going to the terrible resort would not harm her. In other words, our objection is that CCA violates Adequacy Condition 2. Again, we are inclined to think that the ‘preemptive harm’ and ‘preemptive benefit’ problems are serious, but the present objection is independent of them.

As indicated above, the fact that CCA violates Adequacy Condition 2 also means that it has implausible normative consequences. It seems that we always have at least some prudential reason to choose beneficial actions rather than harmful ones. This is captured by the following principle:

Reason for Action. Let possible actions a and a^* be members of the same option-set O_{CS} . If a would benefit S and a^* would harm S , then S has a prudential reason to do a rather than a^* .⁷

Reason for Action is highly plausible. (In fact, we are inclined to accept an even stronger principle, to the effect that the reason to choose a rather than a^* will always be prudentially *decisive*. Here, however, we assume only the weaker principle that there will *exist* such a reason, which need not be decisive.) But CCA and Reason for Action together imply that Beth has a prudential reason to go to the terrible resort rather than to the moderately nice resort. And that is false. Beth has no prudential reason whatsoever to go to the terrible resort rather than to the moderately nice one.⁸

A further problem for CCA is worth noting. Consider again the case with which we introduced the ‘preemptive harm’ problem, where Jill hits Jack with her fist. Again, CCA implies that Jill’s hitting Jack with her fist benefits him, given that she would otherwise have hit him with the baseball bat. Supposing that Jack would have been even better off if Jill had not moved her body at all (and hence not hit Jack), CCA also implies that Jill’s moving her body harms him. The upshot that Jill’s hitting Jack with her fist benefits him, while her moving her body harms him, is odd. Intuitively, it is the other way around: the hitting seems harmful, whereas the moving appears harmless. For example, we might appropriately blame Jill for hitting Jack with her fist, but blaming her for moving her body seems inappropriate.⁹

5. Modifying CCA

Having found strong reasons to reject CCA as well as Absolutism, let us consider well-being counterfactualist theories that try to strike the right balance between those two extremes concerning baselines. It might be thought that some relatively straightforward

⁷ This principle is put forward in Carlson [2019: 797]. As explained there, Reason for Action does not require any particular account of ‘contrastive’ reasons—reasons to perform some action *rather than* some other action. For instance, Reason for Action is compatible with the view that such reasons are reducible to ‘non-contrastive’ ones.

⁸ As explained in Carlson [2019], a similar argument shows that CCA does not capture harm’s *moral* significance. Some writers argue that harm is considerably less morally significant than it might seem (e.g. Feit [2019]). Even if it is, this does not threaten our argument, for the *prudential* significance of harm cannot plausibly be denied.

⁹ Someone who favours a coarse-grained principle of act-individuation might retort that Jill’s moving her body is, in the circumstances, the very same action as her hitting Jack with her fist. Hence, CCA implies that Jill’s moving her body does not harm Jack. However, this reply is problematic. First, on the coarse-grained view, CCA also implies that Jill’s moving her body *benefits* Jack, and this is counterintuitive. Second, the claim that CCA implies that Jill’s moving her body does not harm Jack seems incompatible with the supposition that Jack would have been better off if Jill had not moved her body. Third, identifying Jill’s moving her body with her hitting Jack with her fist implies, given the assumptions of the case, the incoherent-sounding claim that, if Jill had not moved her body, she would have hit Jack with the baseball bat. Although there are replies available to the coarse-grained theorist, these problems indicate that individuating actions coarse-grainedly is not an attractive option in the present context.

modification of CCA could achieve this. A possible lesson to draw from the problems with CCA is that the relevant comparison, for determining whether an event e would harm or benefit a person S , is not what *would* happen to S if e were not to occur, but rather what *could* happen. In order to capture this idea, let us make the following assumption:

Assumption 4. For any possible event e , there is a finite set A of mutually exclusive and jointly exhaustive possible events that includes e and its relevant alternatives. Call such a set an *alternative-set*.

We shall also assume that every option-set is a (proper or improper) subset of an alternative-set. The relevant alternatives to e are the events that could have happened instead of e . If a coin is tossed, for example, the possible event that it lands 'heads' is plausibly an (and perhaps the only) alternative to the possible event that it lands 'tails'. In other cases, we might not always have a clear intuitive idea of how events should be sorted into alternative-sets (in particular concerning possible events that are not actions). But this does not matter for our purposes. We shall eventually argue that no well-being counterfactualist account is acceptable, and our argument does not rely on any particular assumptions about alternative-sets. We can therefore simply assume that there is *some* way for adherents of well-being counterfactualist accounts to sort events into alternative-sets.

We shall now discuss some modifications of CCA that, for each alternative-set A and person S , identify a baseline, W_{SA} , and claim that an event e in A would harm S iff $W_{SA} > W_{Se}$, and would benefit S iff $W_{Se} > W_{SA}$. One such account identifies W_{SA} with the maximum lifetime well-being level for S of the events in A . Let us call this theory *the Best Alternative Account*. Restricted to harm rather than benefit, and to actions rather than events in general, an account very similar to the Best Alternative Account has been proposed by Melinda Roberts [1998: 158–64, 2007, 2009, 2011].¹⁰ Unlike CCA, the Best Alternative Account satisfies Adequacy Condition 2. Problematically, however, the account implies that every alternative except the best one(s), in terms of how well off the respective events would leave S , would harm S . As a result, it fares even worse than CCA with respect to the preemptive benefit problem.

Another problem with the Best Alternative Account is that it appears arbitrary as a general account of harm and benefit. Why not claim, instead, that *only the worst* event in an alternative-set, in terms of how well off the respective events would leave S , would be harmful (cf. Carlson [2019: 802])? That is, why not let W_{SA} be the *minimum* lifetime well-being level for S of the events in A ? The resulting theory, *the Worst Alternative Account*, will not in general yield intuitively less plausible results than the Best Alternative Account does. In situations where all or most alternatives are good, intuitively speaking, the Worst Alternative Account will appear more plausible, whereas the Best Alternative Account will seem more plausible in situations where all or most alternatives intuitively are bad. The Worst Alternative Account fares very poorly with respect to the preemptive harm problem.

Further, the Best and Worst Alternative Accounts both violate Adequacy Condition 1, if construed as accounts of both harm and benefit. The Best Alternative Account absurdly excludes the possibility that an event benefits a person. This can be fixed by weakening the

¹⁰ The account stated in [Roberts 2007] is not strictly well-being counterfactualist, since it includes a causal requirement.

theory's condition for benefit, as follows: e benefits S iff $W_{Se} \geq W_{SA}$. Similarly, the Worst Alternative Account excludes the possibility that an event is harmful. To fix this, we may weaken the theory's condition for harm, to state that e harms S iff $W_{SA} \geq W_{Se}$. Thus revised, however, both theories imply that every event is either harmful or beneficial. Hence, they still violate Adequacy Condition 1, by excluding the possibility of neutral events.

Well-being counterfactualists must thus continue their search for a better version. In particular, they need an account that, unlike Absolutism, CCA, the Best Alternative Account, and the Worst Alternative Account, satisfies both of Adequacy Conditions 1 and 2. One such account, which naturally comes to mind, is *the Average Account*, defined by letting W_{SA} be the average lifetime well-being level for S of the events in A .

Despite having the important advantage of satisfying Adequacy Conditions 1 and 2, however, the Average Account is hardly a satisfactory theory. Suppose that $A = \{e = \text{Gary's just saying 'hello' to Grace, } e^* = \text{Gary's insulting Grace, } e^{**} = \text{Gary's accusing Grace of murder}\}$ is an alternative-set, and that e^* is what actually happens. (Since the events in A are actions, A is also plausibly an option-set.) Suppose also that $W_{Ge} = 10$, $W_{Ge^*} = 8$, and $W_{Ge^{**}} = 3$, with the subscript 'G' standing for 'Grace'. Since the average of these values is 7, the Average Account implies that e^* benefits Grace. Intuitively, however, Gary's insulting Grace harms her. Note, moreover, that the Average Account implies that e^* benefits Grace also if we add the assumption that, if Gary had not insulted Grace, he would just have said 'hello' to her. Thus, e^* comes out as beneficial although it is intuitively harmful *and* Grace would have been better off had it not occurred. Hence, the Average Account arguably fares even worse in this case than CCA does in 'preemptive harm' cases. Any intuitively harmful event that is deemed beneficial by CCA is such that the subject of the putative harm would have been even worse off if the event had not occurred. Hence, there is at least *some* support for the claim that the event is not harmful. As regards Gary's insulting Grace, on the other hand, there seems to be nothing that tells against its being harmful. The mere fact that Gary has an option that would be even worse for Grace does not appear relevant.

6. A General Argument against Well-Being Counterfactualist Accounts

Having rejected Absolutism, CCA, the Best Alternative Account, the Worst Alternative Account, and the Average Account, the prospects for finding an acceptable well-being counterfactualist theory of harm and benefit are beginning to look bleak. In fact, we believe that no well-being counterfactualist account will succeed, and we shall now outline an argument to that effect.

Order the events in an alternative-set A from best to worst, with respect to how well off they would leave a given person S . In the case of ties, any order will do. Let the *difference profile for S of A*, D_{SA} , be the ordered set of differences in well-being levels for S between the events in A , thus ordered. If A has n elements, D_{SA} has $n-1$ elements. For example, if $A = \{e_1, e_2, e_3\}$, $W_{Se1} = 7$, $W_{Se2} = 5$, and $W_{Se3} = 1$, then $D_{SA} = \langle 2, 4 \rangle$. Let *pure comparativism* be the view that whether an event e would be harmful, neutral, or beneficial for S is determined solely by the difference profile for S of the alternative-set containing e . We shall first argue that no version of pure comparativism is acceptable, and then try to show that bringing in absolute well-being levels is of no avail.

Let us start by noting that no difference profile D_{SA} can reasonably be claimed to rule out that the best event in A , with regard to how well off it would leave S , would be beneficial to S .¹¹ Clearly, the best event's being beneficial is compatible with any facts about the differences between this event and other events in A . (Remember that D_{SA} includes no information about absolute well-being levels.) Similarly, no difference profile rules out that the worst event in A , with regard to how well off it would leave S , would be harmful to S . Differently put, for any difference profile for S , there is a possible alternative-set with this profile, such that the best (worst) event for S in this set would be beneficial (harmful) to S . The only way for a purely comparativist account to accommodate these claims is to assume that the best event in an alternative-set is always beneficial, and that the worst event is always harmful.

Since there can be more than two events in an alternative-set, the difference profile must, assuming pure comparativism, determine the status also of alternatives lying between the best and the worst. What about e_2 in the case above, for instance? Since e_2 lies closer to the best alternative, e_1 , than to the worst, e_3 , it might seem natural to conclude that e_2 would be beneficial. If so, it must hold, for all alternative-sets with the difference profile $\langle 2, 4 \rangle$, that the middle alternative would be beneficial. But this is implausible. Suppose that $A = \{\text{Alf's not stealing from Anne, Alf's stealing } \$10,000 \text{ from Anne, Alf's stealing } \$50,000 \text{ from Anne}\}$ is an alternative-set, and that Anne's resulting lifetime well-being levels would be 50, 48, and 44, respectively. Thus, $D_{SA} = \langle 2, 4 \rangle$, implying that Alf's stealing \$10,000 from Anne would benefit her. Intuitively, though, this action would harm Anne, rather than benefit her.

Equally counterintuitive conclusions follow if we instead assume that the middle alternative in all alternative-sets with $D_{SA} = \langle 2, 4 \rangle$ would be harmful or neutral. Consider the set $A^* = \{\text{Alf's giving } \$50,000 \text{ to Anne, Alf's giving } \$10,000 \text{ to Anne, Alf's giving nothing to Anne}\}$, and assume, again, that Anne's resulting well-being levels would be 50, 48, and 44, respectively. Intuitively, Alf's giving \$10,000 to Anne would benefit her, instead of harming her or being neutral. (This verdict is perhaps especially plausible if we suppose that Alf is under no moral or legal obligation to give Anne any money.)

Hence, there are alternative-sets A with $D_{SA} = \langle 2, 4 \rangle$ in which the middle alternative would be harmful to S , and also such sets in which the middle alternative would be beneficial to S . This is incompatible with pure comparativism. And it does not help to assume that every alternative-set has more than three members. The problem that we have highlighted appears also if we consider the status of, for example, the second-best event in an alternative-set with four or more members.

We may thus conclude that harm and benefit are not determined solely by difference profiles. In other words, pure comparativism should be rejected.

Could the solution be to introduce an absolutist element, so as to make harm and benefit a function of difference profiles in conjunction with absolute well-being levels? This is not a promising suggestion. First, the two cases involving Alf and Anne indicate that an event e in an alternative-set A might be beneficial, and another event e^* in another alternative-set A^* harmful, even if A and A^* have the same

¹¹ The only exceptions might be difference profiles like $\langle 0 \rangle$, $\langle 0, 0 \rangle$, $\langle 0, 0, 0 \rangle$, and so on: when A has such a difference profile for S , then a pure comparativist would naturally claim that every event in A would be neutral for S . In the arguments to follow, we set aside difference profiles whose only members are zeroes.

difference profile *and* involve the same absolute levels. If so, bringing in absolute levels obviously does not help.

Second, on pain of violating Adequacy Condition 1, we must avoid the implication that every event resulting in a sufficiently low (high) lifetime well-being level would be harmful (beneficial). For this reason, it is not easy to see what role there is for absolute levels to play. The only remotely plausible possibility, it seems, is to claim that the higher (lower) the absolute well-being level that an event results in, the worse (better) the event has to be in comparison with its alternatives, for it to be harmful (beneficial). Thus, if an event results in a very high (low) absolute well-being level, it is a harm (benefit) only if there is a much better (worse) alternative. This is hardly true in general, however. Suppose that Dan lives a dismal life overall, but that he enjoys eating an ice-cream cone that Dee gives to him. Surely, Dee's giving the ice-cream to Dan can benefit him, even if the resulting lifetime well-being level for Dan is only very slightly higher than the levels in which alternative events would have resulted. Analogously, suppose that Eve has a wonderful life, but that Ernie on one occasion comments critically on her new motorcycle. Surely, Ernie's action might harm Eve, even if no alternative event would have resulted in a much higher lifetime well-being level for Eve.

The upshot is that whether an event *e* would harm or benefit a person *S* does not depend only on (absolute and/or comparative) facts about the lifetime well-being level that *S* would have if *e*, or its alternatives, were to occur. That is, no well-being counterfactualist account of harm and benefit is acceptable.

7. Conclusion

We began by arguing that both Absolutism and CCA face serious problems (sections 3 and 4). We then went on to discuss three more complex versions of well-being counterfactualism—namely, the Best Alternative Account, the Worst Alternative Account, and the Average Account. Unlike Absolutism and CCA, these views allow that whether an event *e* is harmful or beneficial can depend not only on what would have happened but also on what *could* have happened if *e* had not occurred (section 5). Having found those views implausible, we presented a general argument to the effect that the 'difference profile' for *S* of *e*'s alternative-set does not determine whether *e* is harmful or beneficial to *S*, either by itself or in conjunction with facts about absolute well-being levels (section 6). Because this conclusion is incompatible with well-being counterfactualist accounts of harm and benefit, such accounts ought to be rejected.

If well-being counterfactualism is indeed rejected, what view of harm and benefit should we adopt instead? It is natural to think that we could find a more plausible theory by focusing on the *causal*, rather than the merely counterfactual, connections that obtain between an event and an agent's well-being level. However, while such causal theories of harm are popular (see, for example, Harman [2009], Thomson [2011], Gardner [2015], and Purves [2019]), we think that this approach fails too, for reasons that we discuss elsewhere [manuscript]. In our view, there is another possibility that at least deserves serious consideration—namely, that *no* neat and informative theory of harm is true. This is not to imply that the concept of harm should be abandoned, as suggested by Bradley [2012]; after all, many philosophically important concepts lack neat and informative analyses. The point is just that many simple theories of harm and benefit face apparently devastating problems, which tend to remain even

when those theories are complicated in various ways. In view of this, it might be time to question why we should expect there to be any true theory of harm and benefit.¹²

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

Erik Carlson's and Jens Johansson's work for this paper was supported by Grant P14-0212:1 from Riksbankens Jubileumsfond and Grant 2018-01361 from Vetenskapsrådet. Carlson also received support from Vetenskapsrådet's Grant 2016-01531.

References

- Boonin, David 2014. *The Non-Identity Problem and the Ethics of Future People*, Oxford: Oxford University Press.
- Bradley, Ben 2009. *Well-Being and Death*, Oxford: Clarendon Press.
- Bradley, Ben 2012. Doing Away with Harm, *Philosophy and Phenomenological Research* 85/2: 390–412.
- Carlson, Erik 2019. More Problems for the Counterfactual Comparative Account of Harm and Benefit, *Ethical Theory and Moral Practice* 22/4: 795–807.
- Carlson, Erik, Jens Johansson, and Olle Risberg manuscript. Causal Accounts of Harming.
- Feit, Neil 2015. Plural Harm, *Philosophy and Phenomenological Research* 90/2: 361–88.
- Feit, Neil 2016. Comparative Harm, Creation and Death, *Utilitas* 28/2: 136–63.
- Feit, Neil 2019. Harming by Failing to Benefit, *Ethical Theory and Moral Practice* 22/4: 809–23.
- Gardner, Molly 2015. A Harm-Based Solution to the Non-Identity Problem, *Ergo* 2/17: 427–44.
- Hanna, Nathan 2016. Harm: Omission, Preemption, Freedom, *Philosophy and Phenomenological Research* 93/2: 251–73.
- Harman, Elizabeth 2009. Harming as Causing Harm, in *Harming Future Persons: Ethics, Genetics and the Nonidentity Problem*, ed. Melinda A. Roberts and David T. Wasserman, Dordrecht: Springer: 137–54.
- Johansson, Jens forthcoming. Two Arguments for Epicureanism, in *Exploring the Philosophy of Death and Dying*, ed. Michael Cholbi and Travis Timmerman, New York: Routledge.
- Johansson, Jens and Olle Risberg 2019. The Preemption Problem, *Philosophical Studies* 176/2: 351–65.
- Klocksiem, Justin 2012. A Defense of the Counterfactual Comparative Account of Harm, *American Philosophical Quarterly* 49/4: 285–300.
- Norcross, Alastair 2005. Harming in Context, *Philosophical Studies* 123/1–2: 149–73.
- Parfit, Derek 1984. *Reasons and Persons*, Oxford: Clarendon Press.
- Purves, Duncan 2019. Harming as Making Worse Off, *Philosophical Studies* 176/10: 2629–56.
- Roberts, Melinda 1998. *Child versus Childmaker: Future Persons and Present Duties in Ethics and the Law*, Lanham: Rowman and Littlefield.
- Roberts, Melinda 2007. The Non-identity Fallacy: Harm, Probability and Another Look at Parfit's Depletion Example, *Utilitas* 19/3: 267–311.
- Roberts, Melinda 2009. The Nonidentity Problem and the Two Envelope Problem: When Is One Act Better for a Person Than Another? in *Harming Future Persons: Ethics, Genetics and the Nonidentity Problem*, ed. Melinda A. Roberts and David T. Wasserman, Dordrecht: Springer: 201–28.
- Roberts, Melinda A. 2011. The Asymmetry: A Solution, *Theoria* 77/4: 333–67.
- Thomson, Judith Jarvis 2011. More on The Metaphysics of Harm, *Philosophy and Phenomenological Research* 82/2: 436–58.
- Timmerman, Travis 2019. A Dilemma for Epicureanism, *Philosophical Studies* 176/1: 241–57.
- Woodward, James 1986. The Non-Identity Problem, *Ethics* 96/4: 804–31.

¹² Thanks to two anonymous reviewers for this journal, as well as participants in the Higher Seminar in Practical Philosophy at the University of Gothenburg, whose comments led to substantial improvements in this paper.