Harm: The counterfactual comparative account, the omission and pre-emption problems, and well-being

Tanya de Villiers-Botha
Department of Philosophy, University of Stellenbosch, South Africa
tdev@sun.ac.za

The concept of "harm" is ubiquitous in moral theorising, and yet remains poorly defined. Bradley suggests that the counterfactual comparative account of harm is the most plausible account currently available, but also argues that it is fatally flawed, since it falters on the omission and pre-emption problems. Hanna attempts to defend the counterfactual comparative account of harm against both problems. In this paper, I argue that Hanna's defence fails. I also show how his defence highlights the fact that both the omission and the pre-emption problems have the same root cause – the inability of the counterfactual comparative account of harm to allow for our implicit considerations regarding well-being when assessing harm. While its purported neutrality with regard to substantive theories of well-being is one of the reasons that this account is considered to be the most plausible on offer, I will argue that this neutrality is illusory.

Introduction

It has become customary for philosophers interested in defining the concept of "harm" to note that, despite the prominence of injunctions against harming in both ethical theory and common morality, hardly anyone has bothered to say what it is (e.g. Holtug 2002; Bradley 2012; Shiffrin 2012). In an attempt to address this glaring omission, there has recently emerged a relatively small, but rapidly growing, literature that aims to do just that. It is against this background that Bradley (2012) provides a sceptical argument against the likelihood of coming up with a satisfactory definition of harm, citing the problems faced by various attempts thus far (also see Holtug 2002). In the current literature, the definition of harm that has gained the most traction is the counterfactual comparative account, which is considered by many to be the best candidate for a plausible account of harm.[1] Bradley, while agreeing that the counterfactual comparative account of harm is the most promising currently on offer, argues that it is hopelessly flawed, thus strengthening his contention that a plausible definition of harm will remain irrevocably illusive.[2] He argues that the account falters on both the omission problem and the pre-emption problem, both of which he considers to be insurmountable for the theory. He concludes that the concept of harm should simply be dropped from moral discourse altogether.

Hanna (2016) sets out to defend the counterfactual comparative account against Bradley's critique. He attempts to show that, with some modifications, it can overcome both the omission and the pre-emption problems and thus remain the leading contender for a plausible account of harm. I will argue that Hanna fails in his attempts to shore up the counterfactual comparative account against both problems. I will also argue that his defence highlights the fact that these problems both stem from the same limitation, namely that the counterfactual comparative account cannot accommodate the well-being considerations that inevitably play a role in our judgements on the occurrence of harm. Ironically, this claim undermines one of the main reasons why Bradley believes

---

1   See, for example, Feldman (1991); McMahan (2002); Thompson (2011); Klocksiem (2012); Feit (2015); Hanna (2016).
2   For useful overviews of other prominent conceptions of harm, see Hanser (2008; 2013).

that the counterfactual comparative account of harm is the strongest on offer. He describes it as pertaining to extrinsic harm (as opposed to intrinsic harm), meaning that it provides us with a formal definition of harm that is compatible with any substantive theory of well-being.[3] Hence, in addition to concluding that the counterfactual comparative account of harm still faces the omission and pre-emption problems despite Hanna's defence, I will show that it is even weaker than Bradley assumes. I do not necessarily agree with Bradley that the concept of harm should be abandoned altogether; nonetheless, it seems that we will have to look beyond the counterfactual comparative account for a satisfactory conception of harm. A further question that my discussion will highlight is the plausibility of the assumption that a satisfactory definition of harm can be "axiologically neutral" when it comes to conceptions of well-being (Bradley 2012). In other words, questions arise surrounding the plausibility of the intrinsic/extrinsic distinction when it comes to harm. While I do not address this issue in detail here, I point out the possibility that despite the benefits to be had from circumventing the need for a substantive account of well-being in defining harm, it may not be possible to do so.

Bradley's challenge

Essentially, the counterfactual comparative account (CCA) of harm stipulates that an event is harmful when it makes someone (or something) worse off than s/he (or it) would have been had that event not occurred.[4],[5] For example, Robin is harmed by being hit by a bus, seeing that he is worse off for this event than he would have been had it not occurred. As mentioned, the CCA is often touted as the most plausible definition of harm currently on offer. One of its main strengths is said to be the fact that it is an account of extrinsic, overall harm (Bradley 2012; Klocksiem 2012). Since I will partly call this characterisation of the CCA into question, I will briefly elaborate on what is meant with these terms.

Discussions on harm often distinguish between intrinsic and extrinsic harms (e.g. Bradley 2012; Klocksiem 2012; Feit 2015; Hanna 2016).[6] This distinction parallels the distinction commonly made between intrinsic and extrinsic value. Often, in value theory, some things are thought to be valuable in and of themselves, such as pleasure if one is a hedonist. Such things are intrinsically valuable. Other things are thought to be valuable only in as far as they bring about those things that are intrinsically valuable, such as a delicious meal, which brings about pleasure. Such things are extrinsically valuable. Similarly, some things can be thought of as intrinsically harmful (e.g. pain, which is thought to be bad for its own sake), while other things can be thought of as extrinsically harmful (e.g. being hit by a bus, which brings about pain). Bradley (2012, 392) maintains that claims about intrinsic harm are claims about well-being and, as such, are not relevant to "current disputes on the nature of harm". So, the suggestion is that by focusing on extrinsic harm, the CCA can disregard specific theories of well-being.[7] Although Bradley does not elaborate on his grounds for favouring extrinsic accounts of harm, it seems safe to assume that this is due to the difficulties inherent in determining what constitutes well-being.[8] The appeal of any account of harm that is able to accommodate the varying and divergent, existent conceptions of well-being and thus circumvent the need to establish or endorse a specific substantive theory of well-being is obvious. Such an account would make objective evaluations of harm possible. Nevertheless, while I am sympathetic to the attempt to set aside substantive conceptions of well-being in order to define harm, I will argue that the CCA does not, in fact, manage to do so and that this is what lies at the heart of the omission and pre-emption problems. To this effect, I not only argue that Hanna's defence of the CCA against

---

3 Bradley (2012) proposes such "axiological neutrality" as one of the desiderata of a plausible account of harm.

4 There is some debate in the literature as to whether or not definitions of harm should focus on events or states of affairs (Hanser 2008; 2011; Thomson 2011). In this paper, I shall assume that harm pertains to events (as do Hanna and Bradley).

5 Arguably, it is not only people that are subject to harm. For clarity, I only refer to persons in this paper; however, all such references should be taken to include all entities that can plausibly be thought to be liable to harm.

6 Bradley (2012) traces this distinction back to Feinberg (1984).

7 While Hanna (2016) does not explicitly make this claim, he does refer the reader to Bradley's discussion of the merits of the CCA to support his claim that it is an extrinsic account of harm. Later on in his paper, he relies on the intrinsic/extrinsic harm distinction to develop his own definition of *pro tanto* harm (261).

8 For more extensive discussions of these difficulties, see, for example, Moore (1903); Parfit (1984); Scanlon (1998); McMahan (2002).

the omission and pre-emption problems fails, but also that his defence shows why Bradley is mistaken in maintaining that the CCA manages to be axiologically neutral.

The second aspect of the CCA that Bradley considers to be a strength is that it concerns overall harm as opposed to *pro tanto* harm. Very briefly, the idea here is that an event can be prima facie harmful, but turn out to have beneficial consequences overall, which outweigh its harmfulness. For example, an operation can seem harmful in that it entails undergoing an incision and other invasive actions. Yet, if successful, it can leave the patient better off than she would have been had she not undergone it. Hence, the incision and other invasive actions can be considered to be only *pro tanto* harmful, while the operation as a whole is not harmful overall.[9]

With these preliminaries in place, we are now in a position to make sense of Hanna's (2016) formal statement of the CCA, which I shall adopt in this paper:

> CCA: An event *e* is overall harmful for a subject *S* iff *S* is better off in the closest possible world where *e* doesn't occur than *S* is as a result of *e*.[10]

In his critique of the CCA, Bradley (2012) focuses on two problems which he thinks are insurmountable on this account, namely the pre-emption and omission problems. In both of these instances, the CCA seems to get the intuitively wrong result as to the harmfulness of the events in question. Hence, he believes that the CCA should be abandoned altogether, and the concept of "harm" along with it. Hanna (2016) tries to defend the CCA against the omission and pre-emption problems and, in doing so, highlights the root cause of both of these problems.

Hanna's defence of the CCA

In the next two sections, I will address Hanna's defence of the CCA against the omission and pre-emption problems in order to show i) that his defence is unsuccessful, and ii) that his defence highlights a general shortcoming of the CCA.[11]

The omission problem

The omission problem (OP) entails that in omission cases the CCA counts as harmful events that are generally, intuitively, not considered to be so. Usually such cases are construed such that someone fails to receive a benefit in some way. Bradley (2012) uses the following example:

> **Golf Clubs**: Batman buys golf clubs, intending to give them to Robin. That would have made Robin happy. But the Joker convinces Batman to keep the clubs for himself.

The CCA entails that Batman harms Robin by not giving him the golf clubs. This is because, with his omission, Batman leaves Robin worse off than he would have been had Batman, in fact, given him the clubs, given the stipulation that the clubs would have made Robin happy. Bradley's objection here is that "[m]erely failing to benefit someone does not constitute harming that person" (2012, 397). Bradley thus presents *golf clubs* as an instance of failing to benefit someone, which is generally considered to be a problematic type of case for the CCA (see, for example, Shiffrin 2012). The suggestion is that, intuitively, we do not consider simply failing to give someone a benefit to be an instance of harming that person, at least not always. Hence, the CCA overdetermines harms in that it implausibly ascribes harm to such situations.

Hanna undertakes his defence of the CCA against the OP primarily by questioning the *reason* that Bradley gives for the conclusion that a harm has occurred on the CCA in *golf clubs* (viz. that Batman harms Robin because he fails to benefit him). Significantly for our purposes, he does not

---

9   For the purposes of this discussion, I will assume that the *pro tanto*/overall harm distinction as used here is legitimate. I suspect, however, that the distinction rests on the conflation of harm per se and morally permissible harm. It is entirely consistent to hold that the surgeon harms that patient during surgery (i.e. inflicts wounds that will require painful recovery), but that this harming is morally justified, given the patient's consent and/or the greater harm prevented by the surgery. This claim is even more plausible if one were to describe the patient undergoing resultant painful recovery as "suffering a harm".

10 Hanna makes use of the standard possible worlds analysis of counterfactuals, which, as we shall see, in itself highlights some problematic aspects of the CCA. For other notable treatments of the CCA in terms of the standard possible worlds analysis of counterfactuals, see Feldman (1991) and Klocksiem (2012).

11 Where I make use of examples used by both Bradley and Hanna, I use Hanna's construal of the examples.

dispute Bradley's claim that merely failing to benefit someone is not generally considered to be a harm; he disputes the claim that the CCA has this implication. It should be noted that his defence is not aimed at **go⬛ubs** only; it is aimed at all instances of failing to benefit someone. As he result, he makes th⬛ broad arguments: First, he argues that not every instance of failing to benefit someone constitutes a harm on the CCA. This is because the CCA simply does not apply to some failure-to-benefit cases. Secondly, he argues that some failure-to-benefit cases do not come out as harmful on the CCA, as the subject is actually better off as a result of the failure. Finally, he argues that in some instances the finding on the CCA that failing to benefit someone constitutes a harm is correct, although *this is not simply due to the fact that the agent has failed to benefit the subject*. In such cases, there are other factors involved that vindicate the conclusion that harm has occurred. According to Hanna, **go⬛ubs** falls into this last category. I will discuss these three defences separately and show why each fails to shore up the CCA against the OP. I will also discuss how Hanna's defence highlights the strong link between our intuitive ascriptions of harm and concomitant well-being considerations.

Failure-to-benefit does not always entail harm on the CCA
In his first defence, Hanna follows Hanser (2008) in arguing that not every instance of failing to benefit someone constitutes a harm on the CCA, since the definition does not apply to all such cases. If the CCA did entail that all instances of failing to benefit someone constitutes harm, it would mean that Batman harms Robin by not giving him the money in his pocket, by not baking him a cake, by not promoting him to full superhero status, or in any number of other ways in which Batman has it in his power to make Robin better off than he would have been if Batman were to do nothing. This, of course, is precisely the implication that Bradley criticises. Ordinarily, we would not consider someone's failure to give another person the money in his pocket as an instance of him harming that person, even though most people would, *ceteris paribus*, be better off for such a cash injection.[12] Hanna tries to overcome this problem in the first instance by suggesting that the CCA only applies to events and that non-occurrences are not events, meaning that the CCA does not apply to them.[13] To make this argument concrete, consider the following case:

$100: Batman has $100 in his pocket, but does not give it to Robin.

On most, if not a⬛ accounts of well-being, *ceteris paribus*, Robin would have been better off had Batman given him the $100. Hanna's argument is that Robin is not harmed in this scenario on the CCA, given that Batman does nothing whatsoever. That is to say, nothing occurs – there is no event and thus no potential instance of harm. The CCA cannot be applied to a non-occurrence and hence cannot wrongfully entail that Batman harms Robin. Hanna considers this defence to be unproblematic and does not spend any time defending it. The idea is that many apparent failure-to-benefit cases can be eliminated in this way, thus circumventing the OP altogether.

The OP is not so easily overcome, however. To see why, consider the structure of **$100** as I have presented it. In order for the "non-occurrence" defence to work, **$100** needs to portray Batman as entirely passive in relation to Robin's not receiving the money. At no point can he decide against giving Robin the money. In fact, to plausibly constitute a non-event, Batman cannot even *conceive* of the possibility of giving the money to Robin in the first place, as this would leave him in a position where he needs to decide for or against the action. Either that, o⬛ needs to postpone making a decision indefinitely so that he does not perform in a potentially harming event.[14] As a result, the

---

12 A proponent of the CCA can, of course, bite the bullet and argue that any failure to benefit someone is to harm that person. This claim would have the implication that withholding any potential benefit, no matter how insignificant, from anyone, no matter one's relationship to that person, would constitute harming him. The implication is that most of us are harming most people most of the time, as it is almost always the case that we can benefit someone in some way. While this conclusion is not logically faulty, it does detract from the plausibility of the CCA, as, on the one hand, it does not correlate with our intuitive ascriptions of harm and, on the other, does not provide an account of harm that allows for the prudential and normative importance of the concept. Why should we bother attempting to avoid or prevent harm if it is something that is ubiquitous and in⬛⬛ble?

13 Hanna (2016) does not apply this defence to **go⬛bs**, but to failure-to-benefit cases where no mention is made of the potential agent's intentions regarding the benefit at issue.

14 It may be possible to argue that if he were to deliberately postpone making a decision, this in itself constitutes a harming event, as he needs

non-occurrence defence only works in cases where Batman never even considers the possibility of benefiting Robin in the first place. This has the counterintuitive implication that Batman's character plays a major role in determining whether his actions (or non-actions) are harmful on the CCA. The less other-regarding he is, the less likely he is to even consider the possibility of benefiting Robin and the more likely he is to not be involved in a harming event. If he were slightly more altruistic, chances are he might, however briefly, entertain the possibility of benefiting Robin and would thus harm Robin if he were to decide against doing so. The incongruity of this result is further highlighted if we tweak the scenario with regard to Robin's circumstances:

> **Drowning**: Batman comes across Robin who is drowning. It does not occur to him to help. Robin drowns.

On the non-occurrence defence, Batman does not harm Robin in **drowning**, which seems odd. But perhaps we can grant this point. Perhaps our intuition that Batman harms Robin here is due to our moralisation of the scenario and the judgement that Batman has *wronged* Robin by not saving him. This does not necessarily entail that he has *harmed* Robin. If we were to agree that a plausible definition of harm should not commit what Bradley (2012, 395) calls "a moralistic fallacy" (i.e. presupposing the moral principle that harming is wrong), then perhaps we are wrong to insist that Batman harms Robin in **drowning**. We could instead say that it is the drowning event itself that harms Robin and, even though Robin would have been better off had Batman helped him, Batman does not cause the harming event in the first place. So, let us again tweak the scenario to further test the plausibility of the "non-occurrence" defence:

> **Drowning***: Batman unknowingly drops his cape next to a pond. Robin trips over the cape, falls into the pond, and starts to drown. It does not occur to Batman to save Robin. Robin drowns.

On the non-occurrence view, Batman harms Robin by unknowingly dropping his cape next to a pond, but not by failing to fish him out or by not even considering doing so. Nevertheless, we may want to put the incongruity of this result down to the fact that Batman is the one who dropped the cape in the first place, albeit accidently, and hence is responsible for Robin's predicament. The harming event in question here may be the dropping of the cape and not the failure to save Robin. Yet, on the non-occurrence defence, if the Joker had accidently dropped the cape, even while oblivious to the fact that Robin had tripped over it, *he* would have harmed Robin, and Batman would be in the clear. This still seems wrong. Again, the fact that it does not occur to Batman to save Robin is an odd sort of mitigating factor in his culpability for the harm that befalls Robin. If I am right about this, it seems that distinguishing between events and non-events and only characterising events as harmful does not address the omission problem for the CCA in a satisfactory way. It merely serves to exculpate those who have absolutely no regard for others.

Hanna only offers the non-occurrence defence against some failure-to-benefit cases (those where the agent does not consider benefiting someone else). There are other kinds of failure-to-benefit scenarios where the CCA's implications also seem incongruous, and he focuses most of his attention on tackling these. Take a scenario where Batman considers giving Robin the $100 in his wallet, but then decides against it (**$100***).[15] Here, the non-occurrence defence does not apply, and the CCA still entails that Batman harms Robin. So, even if the non-occurrence defence were to be accepted, we would still be faced with the OP in such instances, as it still seems intuitively wrong that Batman harms Robin by not giving him the $100, despite him considering the possibility.

Sometimes the failure to benefit does not leave the subject worse off
In his second line of defence, Hanna tries to save the CCA from the OP by arguing that failing to benefit someone does not always leave that person worse off. Obviously, in such instances, even if

---

to decide to postpone making a decision.

15   Note that **$100*** differs from **golf clubs** in that here Batman never makes the decision to benefit Robin. In **golf clubs** he does make such a decision and then changes his mind. Hanna makes much of this difference.

the potentially harming agent performs an action that constitutes an event, the CCA does not entail that harm has taken place. Consider Hanna's (2016) Batmobile example:

> **Batmobile**: Batman can benefit Robin by giving him the Batmobile. He considers doing so, but decides not to.

On the face of it, the CCA entails that Batman harms Robin. The non-occurrence defence obviously does not apply here, as it is specified that Batman decides not to benefit Robin, so a potentially harming event takes place. Nevertheless, Hanna argues that Robin is not necessarily harmed by setting his sights on the second aspect of the CCA: the requirement that the event in question needs to leave the subject worse off. He casts his argument in terms of the closest possible world account of counterfactuals,[16] echoing points made by Klocksiem (2012). He claims that it is not obviously the case that Robin is harmed, since the closest possible world (CPW) to the one where Batman decides not to give the Batmobile to Robin is not necessarily the one where Batman does decide to do so. In other words, the appropriate counterfactual against which to measure whether or not Robin is worse off is not the one where Batman benefits him (obviously this would have left Robin better off), but one where Batman does something else instead. Hanna is vague on what this something else might be, mentioning that the appropriate CPW might be one where Batman "never decides, say because he remains indecisive or just forgets" (2016, 252).[17] The suggestion is that there are not only two possible options at play here: either Batman gives Robin the Batmobile and benefits him, or he does not and leaves Robin worse off. Other alternatives exist, raising the possibility that Batman leaves Robin better off by not giving him the Batmobile. There are at least two problems that Hanna needs to contend with here: i) plausibility, and ii) the epistemological problem. I will address each separately.

Plausibility
In the first instance, Hanna's defence is simply implausible. He suggests that the relevant CPW with which to compare Robin's situation may be one where Robin is better off for Batman's decision not to benefit him. But it is difficult to see just how this could be. Hanna's own alternatives do not constitute worlds where Robin is better off – Batman's being indecisive or forgetting to decide do not leave Robin better off than he would have been had Batman decided to give him the Batmobile. At best, it could be argued that Robin provisionally breaks even in these CPWs in that Batman does not take a definite decision either way, leaving the possibility open that he may in future decide to benefit Robin. But this defence is inadequate. Firstly, Hanna does not develop an argument as to why these are the appropriate CPWs to use as contrast cases as opposed to the world where Batman decides to benefit Robin. Secondly, these CPWs either have to involve the implausible scenario that Batman never reaches a decision regarding the Batmobile, or a scenario where the potentially harmful event is just postponed to some future date, at which time the CCA would apply and lead to the OP. As such, this defence merely temporarily circumvents the OP by specifying rather unusual circumstances.

Arguably there is no CPW where Robin is unambiguously better off for not having received a benefit other than by stipulating that the "benefit" in question is not a benefit after all. So, Hanna could stipulate that the Batmobile is a death trap, which means that Robin is better off for not receiving it. However, if he were to make this move, he would not be defending the CCA against the OP, as Batman is not failing to benefit Robin but, rather, protecting him from future harm.

Epistemological problem
Even if we were to grant the possibility that there may be a CPW where Robin is better off for not having received a benefit, there is a further problem that is raised by the CPW defence: Just how

---

16   See Stalnaker (1968) and Lewis (2001).

17   Hanna (2016) gives a third option, which is that Batman may never have considered giving Robin the Batmobile in the first place. I take this to be an instance of a backtracking counterfactual (see Lewis 1979), and although Hanna (252, fn. 3) gestures towards defending his use of such a counterfactual in this instance, he does not provide a principled reason for doing so. Hence, I will not address this possibility.

do we identify the appropriate CPW against which to contrast the effect of Batman's decision on Robin? If anything, this defence exacerbates a particularly difficult problem that any counterfactual comparative account struggles with, namely the epistemological problem (EP). As McMahan (2002) notes, the difficulties in determining what would have happened, hypothetically, had a given event not occurred are particularly intractable. While reference to the semantics of closest possible worlds may provide us with a metaphor that is suggestive of how we can go about determining which hypothetical alternatives are the appropriate contrast cases, it does not do much more than that. Practically speaking, how do we determine which is the nearest, relevant possible world, especially if we take into account, as Klocksiem (2012) points out, that there may be no single way in which things would have gone counterfactually? A much more substantial account of how to go about determining what the likely course of events would have been in the absence of the harming event is needed to make this defence work. McMahan (1988; 2002) suggests that the EP may be ameliorated by engaging in evaluations on a statistical or probabilistic basis – i.e. focusing on how things generally tend to go. This suggestion does not help Hanna's case, however, as his defence relies on the very uncertainties that bedevil such probabilistic assumptions. His suggestion is that things may have gone in a direction that we would *not* generally anticipate (Batman reaches no decision, or Robin is better off for not receiving a benefit). So, Hanna is proposing that we consider what we can call non-probable CPWs. While this suggestion may have the effect of saving **golf clubs** against the OP (although it probably does not – see above), it comes at the cost of increasing the array of hypothetical alternative worlds that need to be taken into consideration, without providing an account of how we should go about identifying the appropriate one. If Hanna wants to defend the CCA against the OP in cases such as **Batmobile** by arguing that the subject of the given event may not necessarily be worse off for not being benefited, he needs to develop a plausible account of how to determine how things could have gone for the subject had the agent not failed to benefit him.

Of course, Hanna's non-probable-CPW defence is ineffectual if the scenario at issue is one where the subject would undeniably have been better off had he been benefited. This is the kind of scenario Bradley is alluding to when he raises the OP with regard to **golf clubs** in the first place, and it is at such a scenario that Hanna aims his third defence.

Some instances of failure-to-benefit are indeed harmful
In his third defence, Hanna focuses specifically on Bradley's **golf clubs** scenario and argues that Bradley is right in saying that the CCA implies that Batman harms Robin in this case, but that he is wrong about *why* the CCA entails this. He argues that if this "why" can be correctly characterised, the seeming counterintuitiveness of the conclusion would evaporate, and thus the CCA would be vindicated in this instance. Hanna puts the harm that occurs in **golf clubs** down to an additional feature of the case, in contrast with **$100** and **Batmobile** – here, Batman *intends* to give Robin the golf clubs but *changes his mind*. **Golf clubs** is clearly subject to the CCA in that there a potentially harming event takes place – Batman decides not to benefit Robin. In addition, the **Batmobile** defence is not an option, as the relevant CPW is limited to the world in which Batman does not change his mind. Normally, on the CCA (assuming that the golf clubs are a benefit, which seems unproblematic since no reason is given to think they are not), Robin is considered to be worse off for not receiving them, and thus Batman harms him. This is the conclusion that Bradley finds counterintuitive. Hanna counters Bradley by arguing that the reason that **golf clubs** comes out as harmful on the CCA is *not* due to the fact that Batman fails to give Robin the golf clubs, but due the fact that Batman changes his mind.

In his defence, Hanna focuses on the specification given by Bradley in the original scenario that Batman intends to give Robin the Batmobile, but then changes his mind due to the Joker's influence. He argues that, it is clear that the Joker harms Robin *by preventing him from getting a benefit*.[18] So, the Joker's actions clearly constitute a harm on the CCA, as he makes Robin worse off than he otherwise would have been. Problematically, Hanna then argues that if we accept that the Joker's

18　When a third party prevents a subject from getting a benefit, it is generally considered to be a genuine instance of harm (see, for example, Hanser 2008; Shiffrin 2012).

actions harm Robin, we have to accept that Batman's actions do so as well. The idea is that the two actions are inseparable in terms of their contribution to Robin's state – Batman only changes his mind because the Joker convinces him to. And since we have already established that the Joker's action harms Robin, by extension, Batman's does as well. Contra Bradley, Hanna sees this result as correct. However, he does not see it as entailing that *all* failures-to-benefit are harms under the CCA. He sees this case as significantly different from cases such as **$100** and **Batmobile** in that there are two additional factors at play: an additional agent who becomes the primary instigator of the harm, and the stipulation that Batman had initially intended to give the golf clubs to Robin. Nevertheless, neither of these additional factors have the redeeming implications for the CCA that Hanna takes them to have.

The problem for Hanna's overall argument is that the Joker does not harm Robin by omission, which, on his argument, means that Batman does not either (since Batman's actions merely facilitate the Joker's). So, in this defence, Hanna obscures the core feature of the scenario that Bradley objects to: the fact that CCA there entails that the failure to benefit (i.e. an omission) is harmful. This is the feature that makes **golf clubs** relevantly similar to **$100** and **Batmobile**. Strictly speaking, the Joker is incidental to the case. It could just as well be set up as follows without losing any of the relevant features that Bradley objects to:

> **Golf Clubs***: Batman buys golf clubs with the intention of giving them to Robin. He changes his mind and keeps the clubs for himself.

On the CCA, Batman harms Robin. He fails to give Robin a benefit, making him worse off than he would otherwise have been. This is what Bradley finds counterintuitive. Whether or not Batman changes his mind due to the Joker's influence is neither here nor there. Hanna uses the Joker's presence in the scenario to obscure the fact that this is a straightforward failure-to-benefit case as far as Batman's actions go, which means that Hanna does not defend CCA against the OP. Nor is the fact that Batman changes his mind significant in terms of the counterintuitiveness of the result. For this aspect of the scenario to matter, Hanna needs to develop an argument as to why it is relevant that Batman changes his mind, which he does not do. In addition, by attributing the harm occurring in this scenario to Batman's changing his mind again serves to obscure the fact that Bradley's criticism is aimed at straightforward failure-to-benefit cases. It is against such cases that the CCA needs to be defended in order to overcome the OP. Hanna does not defend the CCA against such a case here, meaning that Bradley's criticism still stands.

In the above, I have addressed the three ways in which Hanna attempts to defend the CCA against Bradley's contention that it falters on the OP. In each instance, I showed that Hanna does not succeed. In order to counter the OP, Hanna either needs to show that the CCA is right in entailing that Batman harms Robin in all cases where he fails to benefit him, and that it is our intuitions that fail if we think otherwise, or he needs to find a principled way in which to distinguish between failure-to-benefit cases that plausibly entail harm and those that do not. Hanna does neither, which means that the CCA is still subject to the OP.

In my introduction, I claimed not only that Hanna's defence of the CCA against the OP (and the PP) fails, but also that it highlights the role that well-being judgements implicitly play in our ascriptions of harm. I put the significance of this claim down to the fact that the CCA is often presented as pertaining to extrinsic harm and thus as being able to accommodate the different conceptions of well-being that may inform harm judgements. I will discuss this aspect of argument in more detail below. For now, it is instructive to note that Hanna does not call Bradley's contention that failing to benefit someone is not (necessarily)[19] to harm them into question. Instead, he attempts to circumvent the OP by either denying that the CCA entails that failure to benefit is a harm or by arguing that when the CCA does entail this, the harm done is due to something other than merely failing to benefit someone. My contention is that it is necessary for him to try and evade the

---

19   Bradley's (2012) critique seems to suggests that failing to benefit someone never constitutes a harm. This is implausible, as, intuitively, there are cases where failure to benefit someone may well be to harm them, such as when someone well off refuses a starving person a life-saving meal.

omission problem in this manner, because it cannot be addressed by the CCA as it stands. This is because it cannot accommodate the role that our implicit conceptions of well-being plays in our intuitive assessment of whether or not harm occurs when someone fails to benefit someone else.

The pre-emption problem

In the introduction, we saw that Bradley believes that there are two fatal problems for the CCA, namely the omission problem (OP), as discussed above, and the pre-emption problem (PP). Hanna also aims to defend the CCA against the PP. As with the OP, my argument will be that he fails to do so.

Pre-emption cases are structured so that an intuitively harmful event befalls a subject moments before a second, intuitively equally harmful event. On the CCA, incongruously, the first event is rendered not-harmful, due to the existence of the second.[20] Bradley (2012) uses the following case to illustrate the problem:

> **Heart attack**: Batman dies of a heart attack. A millisecond later he is hit by a flaming cannon ball that would certainly have killed him.

On the CCA, the heart attack does not harm Batman, as he is not worse off for it than he would have been had it not occurred, given that he would have been killed moments later by the flaming cannon ball anyway. Intuitively, this seems to be the wrong result. Generally, dying from a heart attack is considered to be a harm.

One option for proponents of the CCA is to capitulate and accept that the heart attack does not harm Batman. Many (including Bradley) find this implausible and hold that the PP poses an insurmountable problem for CCA (e.g. Hanser 2008; Shiffrin 2012). The objection is that this result runs counter to our ordinary ascriptions of harm – generally, heart attacks are considered to be harms whether or not they pre-empt alternative deaths. Hanna attempts to redeem the CCA and preserve the intuition that the heart attack is harmful by arguing that the heart attack is a *pro tanto* harm, but not an overall harm (see my discussion above).[21] I will argue that Hanna's *pro tanto* defence fails and that the PP remains an obstacle for the CCA.

Pre-emption and pro tanto harm

As discussed above, proponents of the CCA often distinguish between *pro tanto* harm and overall harm as a way to account for cases such as **operation**, where a patient is prima facie harmed by a surgeon, but where she is better off overall for having undergone the operation. Hanna calls upon this distinction to defend the CCA against the PP. Applying it to **heart attack**, he argues that the *pro tanto*/overall harm distinction has the following implications: The heart attack is *pro tanto* harmful for Batman if it brings about something intrinsically bad for him (such as pain). On the other hand, it is not overall harmful for Batman, if it ends up having more beneficial consequences than *pro tanto* harmful ones (such as saving him from the greater pain of being hit by the cannon ball). So, by applying this distinction to pre-emption cases, we can account for the intuition that the heart attack is somehow harmful and also for the fact that its harmfulness is affected by the second potentially harmful event, viz. the cannon ball hurtling towards Batman.

Yet, Hanna's deployment of the *pro tanto* defence is not legitimate here. Recall that the *pro tanto*/overall harm distinction is generally employed to account for the intuition that some harms are not necessarily bad, such as when a rescuer needs to break a patient's arm in order to save her from drowning (**rescue**) (see Shiffrin 1999; 2012). But there is an important difference between pre-emption cases and cases such as **operation** and **rescue**. In both, there is a clear causal link between the *pro tanto* harming event and the potentially worse harming event that it prevents.

---

20   Usually in these thought experiments there is a minimal time interval between the two events. But, it is not obvious that this needs to be the case. McMahan (1988; 2002) presents variations on such cases where the time interval between the potentially harmful events is varied. This highlights another potential problem for this line of defence. Arguably, many harms befall all of us at varying intervals over the course of our lifetimes. What distinguishes cases such as **heart attack** (see below) from such run-of-the-mill cases of multiple harms? What length of time between harms is required for them to be considered to be unrelated and hence not subject to the PP?

21   Klocksiem (2012) makes a similar argument.

The *pro tanto* event is unambiguously necessary to prevent the worse harm from occurring. In **rescue**, for example, it is stipulated that the rescuer needs to break the patient's arm in order to save her; the harm of the broken harm prevents the greater harm of drowning. Pre-emption cases do not necessarily have this feature. The harming events follow upon one another in relatively close proximity, but are otherwise contingent (it remains an open question as to what kind of time lapse between two harming events would disqualify two such events from constituting a case of pre-emption). This is what drives the prevalent intuition that the CCA gets it wrong when it entails that the heart attack is not harmful to Batman, since he would have died by flaming cannon ball moments later. The heart attack happens to save Batman from a fiery death, but it brings him no other advantage, nor is it necessarily a lesser harm than being struck by a cannon ball. Batman still ends up dead.

Hanna may want to argue that **heart attack** is underdescribed, as he implicitly does when he stipulates that death-by-fiery-cannon-ball would have been more painful than death-by-heart-attack. This makes his claim of *pro tanto* harm more plausible, as Batman is saved from the ostensibly greater harm of dying painfully. Yet, this move does not save the CCA from the PP. Firstly, one may question whether the conclusion that Batman merely suffers a *pro tanto* harm, but not an overall harm from the heart attack truly does track our intuitions. It seems absurd to hold that Batman would have suffered an overall harm if he were to die from a heart attack in the absence of the looming cannon ball, but that he is actually benefited by the heart attack, given the impending cannon ball (and the pain that it would cause). Instead, the two scenarios seem to be on a par – dying from a heart attack is bad, all things considered, and so is dying from being struck by a flaming cannon ball. In addition, the *pro tanto*/overall harm distinction would not be applicable to pre-emption cases where the two events at stake are equally intrinsically harmful (e.g. equally painful). If the heart attack and the canon ball were equally painful, the PP would arise again, as Batman is not better off than he otherwise would have been on suffering the heart attack. Even in the instance where the first event is intrinsically more harmful than the second event, the reason why it comes out as harmful on this defence of the CCA seems wrong. Now, Batman is harmed by dying from the painful heart attack as opposed to dying from the less painful cannon ball. This still does not track the intuition that simply by dying Batman suffers a harm, irrespective of what would have happened moments later.

It is open to Hanna to argue that it is our intuitions in this instance that are wrong, especially given that the harm of death is notoriously difficult to accommodate on most accounts of harm.[22] However, another variation on the pre-emption case that does not involve death also grapples with the PP.

Free acts, intentions, and the PP

If Hanna's modified *pro tanto* defence as discussed above seems plausible with regard to **heart attack**, it seems less so in a variation on the pre-emption problem that Bradley (2012) also discusses, namely **Bobby Knight**:[23]

> **Bobby Knight**: Bobby Knight chokes Philosopher who says "hi" to him, thereby crushing and permanently damaging her windpipe. If he had not been attending anger management classes, he would have dismembered her instead.

On the CCA, Philosopher is better off for having a permanently damaged windpipe than she would otherwise have been – if Knight had not choked her, he would have dismembered her. Hence, Philosopher is not harmed. Yet, suffering permanent damage to one's windpipe is generally considered to be a harm. The *pro tanto* defence is ostensibly again open to the proponent of CCA – the damaged windpipe is a *pro tanto* harm, but not an overall-harm, as it saves Philosopher from

---

22   See McMahan (2002) for an extended discussion.
23   This variation on the pre-emption problem is due to Norcross (2005), who attempts to demonstrate a shortcoming in the CCA which we have already discussed, namely the failure to stipulate a principled way in which to identify the relevant CPW to serve as contrast point to a given event.

dismemberment. However, the upshot of this defence is that Knight could be said to have benefited Philosopher by choking her (assuming that benefit and harm are treated analogously here[24]), which many find to be an implausible result. Bradley (2012) suggests that describing the harm here as *pro tanto* underplays the harmfulness in question, stating that serious overall harm occurs as well. Certainly, while it may very well be that things could have gone worse for Philosopher, describing her as unharmed overall after her encounter with Bobby Knight is highly counterintuitive. As Bradley (2012, 407) puts it, this case seems to imply that "one can make one's pain-causing actions overall beneficial rather than harmful merely by becoming a rage-filled lunatic who would otherwise have done much worse things".

Hanna deals with **Knight** in two ways: Firstly, he focuses on Bobby's degree of control over his behaviour. He argues that if the only way in which Bobby could have prevented himself from hurting Philosopher more severely was to limit himself to crushing her windpipe, then his behaviour can legitimately be described as overall beneficial to her, and hence the CCA is vindicated. He does not offer much by way of argument here, and there is little for someone who has the opposite intuition about the case to engage with. Hanna's defence appears to be informed by a moralistic fallacy, as discussed above. He seems to suggest that the fact that it may be morally permissible to excuse Bobby's behaviour given his psychological state affects whether or not he has harmed Philosopher. But such considerations should not influence our attributions of harm, which are supposed to precede any moralisation of the phenomenon. If Philosopher were to trip and fall and so damage her windpipe, we shall have no trouble in saying that she suffered a harm, even though it remains entirely possible that she may have been injured more severely in her fall. In terms of a plausible definition of harm, whether or not the damage to Philosopher's windpipe is caused by a fall or by Knight should make no difference. Only once it has been established that Philosopher has, in fact, been harmed, does the moral culpability of the agent involved (if there is one) become relevant. Hence, Hanna's first stab at defending the CCA against **Bobby Knight** fails.

In his second line of defence, Hanna focuses on what he believes to be the real problem for the CCA in the **Knight** case, namely, the possibility that Knight could have stopped himself from choking Philosopher, but does not do so. He goes on to address the counterintuitive result that Knight does not harm Philosopher by arguing that we should be using a different possible world than the one usually used (i.e. the world where Knight dismembers Philosopher) as a point of comparison in this case, namely the one where Knight does not assault her at all.[25] Hanna justifies this specification on two grounds: 1) Both acts are associated with the same *intention* on the part of Bobby (i.e. hurting Philosopher) and so should be treated as a collective; and 2) Bobby *could* have refrained from hurting Philosopher altogether, given Hanna's stipulation that his action was free (presumably the idea here is that if his action were free, his refraining from it was a live option).

So in this second defence, Hanna places a great deal of emphasis on the consideration of whether or not Knight could have acted otherwise. In fact, Hanna develops a revised principle (CCA+) to deal with free acts specifically (2016, 266):

> CCA+: An agent's free act is overall harmful for a subject *S* iff *S* is better off in the closest possible world where the agent doesn't freely perform an act of that type than *S* is as a result of the act.

Hanna's justification for developing CCA+, namely that Bobby could have refrained from hurting Philosopher, is subject to a similar criticism to that given above regarding Knight's ability to control his actions. In stipulating that Knight's actions are free, Hanna is simply reiterating that he had control over what he was doing and could have done otherwise. In treating free acts differently from mere events, Hanna is presupposing a moralised conception of harm. He is specifying that where

---

24 See Shiffrin (2012) for a critique of the CCA and other comparative definitions of harm for treating harms and benefits symmetrically.

25 Norcross (2005) does something similar in that he argues that the choice of appropriate CPW to use as a comparison case is context sensitive. Briefly, according to him, in the case of Knight, the appropriate CPW with which to compare Knight's actions is not the one where he tears Philosopher's arms off, but the one where he simply says "hi" and walks on, since this is what would be likely for most other people. However, Norcross does not develop an argument for why the world which is likely for most other people is the appropriate CPW in this instance.

agents could have acted otherwise, actions that leave subjects worse off are harmful; whereas, if the agent could not have acted otherwise, the actions are not harmful. Two issues are being conflated here: harmful actions and culpable actions. But, as argued above, culpability should not factor into a plausible definition of harm, which should not presuppose a moral theory. In addition, the only reason for developing the CCA+ and singling out free acts is to shore the CCA up against the PP as set out in **Bobby Knight**, which makes this defence unacceptably ad hoc. There is no other advantage to the CCA in making the distinction between events and free acts, and any other theory that is able to accommodate the Knight case without such ad hoc amendments would trump the CCA. Over and above these problems, proposing the CCA+ in itself is not enough to defend the CCA against the Knight PP. Hanna's defence against the PP in **Knight** also has him appealing to Knight's intentions in order to justify treating both the choking and the dismembering actions as a collective, which allows him to argue that the appropriate CPW to the one where Knight chokes Philosopher is not the one where he dismembers her, but the one where he does not assault her at all. This is to pre-empt the objection that the CCA+ still entails that Philosopher is better off for being (freely) choked if Knight could (freely) have dismembered her instead. This appeal to intentions has its own set of problems.

With regard to treating Knight's actions as a collective due to their being driven by the same intention, Hanna faces the obvious difficulty of specifying how to establish what Knight's actual and most relevant intentions are.[26] But even if it were possible to establish Knight's intentions with relative ease, the appeal to intentions is again ad hoc. As it stands, on Hanna's theory, the intentions of the agent only come into play in possible pre-emption cases, specifically ones where free acts are concerned. But if the intentions behind the agent's actions affect their harmfulness in such instances, would they not do so in other instances as well? Even more problematically, Hanna assumes that Knight has an identifiable fundamental intention, but gives no reason as to why this should be so. The possibility remains that Knight was, at base, motivated by conflicting intentions, which resurrects the possibility that he, in fact, benefited Philosopher in that the CPW to the one in which he damages her windpipe is the one in which he dismembers her.

Over and above these problems, Hanna's defence of the CCA in the **Bobby Knight** case does not even address the substance of Bradley's criticism. The substance of Bradley's criticism is that the result on the CCA that Philosopher's permanently damaged windpipe does not constitute a harm is counterintuitive. Whether or not the harm was caused by a free agent is irrelevant. As with the OP, we can cast the pre-emption problem in different terms, where Bradley's criticism still applies, but where Hanna's defence is ineffective:

> **Bobby Knight***: Philosopher trips and falls over Bobby Knight's discarded book, thereby permanently damaging her windpipe. If she had not tripped, she would have been dismembered by the out-of-control lawnmower approaching her.

The essence of Bradley's criticism still applies to this example. It seems wrong to conclude that Philosopher has not been harmed by permanently damaging her windpipe even though tripping over the book left her better off than she otherwise would have been. Hanna's free-act defence does not apply to this case, so the only recourse open to him is to argue that Philosopher suffered only *pro tanto*, but not overall harm. But this implies that a permanently damaged windpipe is not harmful in an instance where worse harm could have befallen the patient, but harmful when no such alternative harm was an immediate prospect. Ordinarily, we would consider having a permanently damaged windpipe to be a harm. The fact that in incurring the harm, Philosopher avoided alternative, worse harm does not leave her any less harmed. We may want to comfort her in her misfortune by pointing out that she was lucky not to have suffered a worse fate, but it is doubtful whether she would consider her damaged windpipe not to be a harm because things may have been worse. Counterfactually, she could have suffered other harms, but this possibility does not negate the fact that she has been harmed. This consideration highlights the weakness inherent in the *pro tanto* defence when it comes to pre-emption cases. While it is plausible to hold that the necessary harm suffered to facilitate an

---

26   See Hanna's (2016) extended discussion on more fundamental intentions and motivational priority, and so forth.

operation is off-set by the overall benefits resulting from the operation, it is much less plausible to hold that a serious injury is somehow not a harm simply because, contingently, things could have been worse. While the *pro tanto* harm inflicted during the operation is a necessary prerequisite to the performance of the operation, which justifies the claim that the harm and benefits resulting from this situation are interrelated, there is no necessary link between the events in pre-emption cases. They merely happen to follow shortly upon one another, and neither occur in service of a greater benefit. Any benefit to be had here is entirely incidental. This explains why it seems more plausible to describe the operation as *pro tanto* harmful, but not the heart attack or Philosopher's damaged windpipe.

In this section I have argued that Hanna does not manage to defend the CCA against the PP, as presented by Bradley. Not only does the analogy between **heart attack** and **operation** not hold, which undermines his *pro tanto*/overall harm defence, he arbitrarily distinguishes between events and free acts and also appeals to Knight's intentions to preserve the plausible intuition that harm occurs in **Bobby Knight**. In addition, *the pro tanto*/overall harm defence as applied to **Bobby Knight** is subject to the same criticism as in **heart attack**.

As with Hanna's treatment of the omission cases, he does not manage to address pre-emption cases in their most straightforward form and his discussion ends up obfuscating what lies at the heart of the problem. In the final section of this paper, I will elaborate on why I believe that the OP and PP stem from the same aspect of the CCA, namely that it focuses on only one of the factors that we tend to take into account when ascribing harm, namely how things would probably have gone had the event in question not occurred. Yet, this is not the only, or even necessarily the primary, consideration that goes into our intuitive judgements. My contention is that the OP and PP are the result of an additional dimension (or, *dimensions*) of comparison that we tend to employ when ascribing harm. As a result, they will remain intractable under the CCA.

The CCA and well-being
In this section, I argue that the OP and PP show that the CCA is, at best, incomplete as a plausible definition of harm and that the above discussion highlights the fact that the OP and PP stem from the same root cause – the role that our implicit judgements regarding (actual and/or ideal) well-being play when ascribing harm.[27]

The CCA establishes harm simply by determining whether the subject to a given event is made worse off than she would have been had the event not occurred, irrespective of her levels of well-being before, at, or after the event. As a result, since our intuitive conceptions of harm are necessarily intertwined with such well-being considerations, the CCA inevitably diverges from these conceptions in some cases. Arguably, in run-of-the-mill cases, the role played by well-being judgements is generally obscured, and the CCA tracks our intuitions. However, omission and pre-emption cases are such that well-being considerations come to the fore, causing the CCA to diverge from our intuitions.

Clearly, when determining whether someone is better or worse off for a given event, one needs some baseline against which to measure the change in his or her circumstances that the event causes. My contention is that this baseline is generally not what would have happened had the change not occurred. Granted, it often is the case that when someone is left worse off than they might have been, we consider that person to be harmed. However, what primarily drives our harm judgements in these cases is that a person's well-being has been affected in an appropriate way.[28] (I suspect that this could pertain both to someone's actual or their ideal well-being, which would account for harm ascriptions in cases where someone's actual well-being is not necessarily impacted, such as with Parfit's (1984) non-identity cases). The apparent success of the CCA is due to the fact that the role played by such considerations is often obscured in that both the CCA-based

---

27   See Holtug's (2002) discussion of what he deems to be the main theories of welfare and their limitations with regard to defining harm.

28   I am setting aside questions regarding the plausibility and merits of well-being-based definitions of harm for now, as my primary aim here is to show how well-being-based ascriptions of harm can dovetail with CCA-based ones. For overviews of problems that such accounts of harm face, see Holtug (2002); Hanser (2008); Thomson (2011); Bradley (2012).

and well-being-based considerations generally dovetail and lead to similar ascriptions of harm, especially when the cases involved are underdescribed. In the scenario where Robin is hit by a bus and maimed (with no additional information given), the CCA and most theories of well-being entail that he has been harmed. On the CCA – assuming that Robin experiences sufficient well-being – Robin is worse off than he would have been had the bus not hit him.[29] On most theories of well-being, being maimed is inherently bad, whether it be due to the pain caused, the loss of bodily integrity, or the loss of a number of other goods, etc. Here, the CCA tracks intuitions based on well-being considerations, and seems plausible. The problem comes in where well-being is not sufficiently affected by an event that leaves a subject worse off (leading to the OP) or, alternatively, where well-being is obviously affected, but this cannot be accommodated by the CCA (leading to the PP). In such instances, the limitations of the CCA become apparent. Let us look at the OP first.

Well-being and the OP

The contention is that the omission problem arises when a subject's well-being is not obviously negatively affected (or not sufficiently affected) by an event, but where the CCA finds that she has been harmed.[30] In underdescribed cases – where no explicit information regarding well-being is given – our implicit assumptions concerning the subject's (and perhaps the agent's) well-being will play a major role in our intuitions as to whether or not harm has occurred. Take an ordinary omission case, such as **$100**. In the absence of any reason not to, we tend to make the default assumption that Robin is experiencing sufficient levels of well-being and that he is not in dire need of Batman's $100.[31] That is why it seems wrong to conclude, on the CCA, that Batman harms him by not benefiting him. If we knew that Robin was in dire need of the money, this conclusion would seem more plausible. Applied to **golf clubs**, we can see that the reason that the CCA result is counterintuitive is not primarily due to whether or not an event takes place, or whether or not Batman reneges on an explicit intention, as Hanna suggests. It has to do with our implicit assumptions regarding the extent to which Robin needs the clubs (and even the extent to which Batman does not). Bradley's intuition that Batman does not really harm Robin by not giving him the golf clubs is due to his (plausible) implicit assumption that the golf clubs will have a relatively negligible impact on Robin's well-being, since no reason is given to suppose that Robin is badly off in any relevant way that leaves him in need of the clubs. If we were to specify (**golf clubs***) that Robin has been diagnosed with clinical depression and that his psychiatrist has suggested that he take up golf, but that he cannot afford a set of clubs, our intuitions regarding the harmfulness of Batman's actions may come out differently. Here, it is more plausible to conclude that Batman harms Robin, especially if we were also to stipulate that Batman has no great need of the golf clubs. Our intuitions are driven by considering the levels of well-being of those involved and not whether or not Robin would be worse off for not having the clubs per se. If the clubs are taken to be a benefit *simpliciter*, Robin *is* worse off for not having the clubs, regardless of his level of well-being. However, it is only on some theories of well-being – those where his well-being is affected in a relevant way – where his being worse off is considered to be a harm. The fact that tweaking the scenario, so as to specify that Robin's (and Batman's) levels of well-being affects our intuitive ascriptions of harm, shows that it is relative well-being that our intuitions latch on to and not whether or not he is worse off as such.

The upshot of the above discussion is that the CCA only tracks our general intuitions regarding harm when it correlates with our intuitions about the relative well-being of the subjects involved. As soon as the CCA and our well-being judgements diverge, the CCA comes out wrong. In omission cases, it becomes clear that the subject simply being worse off for an event is not sufficient reason to conclude that he has been harmed. His well-being needs to be affected in the appropriate way for that to be the case. This lies at the heart of the OP.

---

29   If I were to specify that Robin was suffering from a painful, terminal illness, the outcome on the CCA could possibly be different.

30   The extent to which the subject's well-being needs to be negatively affected for harm to occur will differ depending on the theory of well-being employed, but all such theories will have some threshold below which the subject will be considered to be harmed.

31   I will not go into detail regarding possible reasons for this assumption. Suffice it to say that it may have to do with what is most familiar to readers (i.e. readers generalise from their own case), or with what comic book fans know about Robin. Given the level of detail given, we have no reason to suppose that Robin is badly off.

The pre-emption problem and well-being

In pre-emption cases, the CCA finds that someone undergoing an event that is usually considered to be harmful is not harmed, due to a second ostensibly harmful event that follows upon the first, as is illustrated in **heart attack**. The pre-emption problem follows from the implausibility of this result, as the harmfulness of the first event should not be affected by the looming second event. Again, my argument is that the divergence between the CCA's findings and our intuitive ascriptions of harm is due to the fact that our intuitions are not based simply on whether or not the subject is worse off, but on whether or not his well-being has been affected in an appropriate way. This is why, in **heart attack**, the fact that Batman would have died by fiery cannon ball anyway does not seem to justify the result that dying from the heart attack is not harmful. Intuitively, heart attacks are harmful. In the absence of a reason to think otherwise, we assume that Batman experiences sufficient levels of well-being and that death negatively impacts on that well-being. That is what makes his death harmful. That Batman would have died seconds later in any event does not make his death any less harmful. This is also why Hanna's *pro tanto* defence fails to convince. For it to work, Hanna has to stipulate that death-by-cannon-ball would have been more painful than death by heart attack. This is supposed to entail that the heart attack is actually a blessing in disguise and thus not a harm. However, if we take the effect of the heart attack on Batman's well-being into account, we find that Batman is not better off on the whole for having avoided a fiery, perhaps more painful, death. He is still dead. He experiences no well-being whatsoever and has thus suffered the ultimate harm. In contrast, if we were to specify that Batman was in agony before his death and preferred to die, or that he preferred death-by-heart-attack to death-by-cannon-ball, we could plausibly conclude that his death did not come as a harm. This has nothing to do with whether or not he would have died in any case, but with whether his well-being was such that his specific death could be considered to be a better alternative than living.

This point is perhaps more clearly illustrated in the **Bobby Knight** case. The conclusion on the CCA that Philosopher is not harmed seems wrong, not because Knight could have acted otherwise or because he had the fundamental intention of hurting her, but because having a permanently damaged windpipe negatively impacts one's welfare on just about any plausible conception of well-being. Whether or not Philosopher could have been hurt more badly in the incident does not detract from the fact that her well-being has been significantly diminished. That she may have suffered alternative, possibly more severe, harms may lead us to conclude that she was lucky to have escaped a worse fate, but that does not make the conclusion that she has not been harmed or, worse, that she was benefited by Knight, plausible.

My overall point here was to show that our intuitions regarding the harmfulness of actions are driven by welfare considerations rather than by simply considering whether or not subjects are made worse off by events than they might otherwise have been. The OP and the PP arise when the CCA and well-being considerations diverge, thereby highlighting the fact that our intuitions track something other than being counterfactually worse off when ascribing harm. Of course, it remains open to the proponent of the CCA to argue that it is our intuitions that are in the wrong in such instances. In order to preserve the CCA as a plausible theory of harm, such theorists would then need to find principled ways in which to overcome the OP and PP. I have argued that Hanna does not manage to do this, and I suspect that it cannot be done while preserving the CCA's neutrality regarding theories of well-being.

Conclusion

Intuitively, we do not consider all events that leave someone worse off to be harms. If we did, all failures to benefit would be harms, which seems absurd. I have argued that it is in fact the extent to which someone's well-being is negatively affected by an event that primarily determines whether or not we consider him to be harmed. Nor do we only consider those events that leave us worse off overall to be harms. If we did, the conclusion that Philosopher is unharmed despite having a crushed windpipe would not seem odd. My argument here was that her well-being was negatively affected in an appropriate way, and this is why it seems wrong to conclude that she is unharmed, despite not being worse off than she would have been counterfactually. Such cases highlight the implicit

role that well-being considerations play in our attribution of harm. The inability of the CCA to factor in such well-being considerations means that it fails as a satisfactory account of harm.

Despite this conclusion, I do not agree with Bradley that the concept of harm needs to be dropped from the moral lexicon altogether. There may very well be other plausible conceptions of harm that are not subject to the problems discussed here. Pragmatically speaking, given the prevalence of harm considerations in our moral theorising and everyday moral judgements, dropping the concept seems unfeasible. In addition, to my mind, the avoidance of harm carries with it the kind of prudential importance that justifies the centrality of the concept to morality. Given the arguments made above, I contend that any plausible conception of harm will have to allow for well-being considerations in some way, especially if it wants to accord with our general intuitions on harm. Ideally, such a definition should remain neutral regarding substantive theories of well-being, given the intractable nature of the disputes around the concept. Whether this is at all possible, remains to be seen.

References

Bradley, B. 2012. "Doing Away with Harm." *Philosophy and Phenomenological Research* 85 (2): 390–412. doi:10.1111/j.1933-1592.2012.00615.x.

Feinberg, J. 1984. *The Moral Limits of the Criminal Law: Harm to Others*. Oxford: Oxford University Press.

Feit, N. 2015. "Plural Harm." *Philosophy and Phenomenological Research* 90 (2): 361–388. doi:10.1111/phpr.12033.

Feldman, F. 1991. "Some Puzzles About the Evil of Death." *Philosophical Review* 100 (2): 205–227. doi:10.2307/2185300.

Hanna, N. 2016. "Harm: Omission, Preemption, Freedom." *Philosophy and Phenomenological Research* 93 (2): 251–273. doi:10.1111/phpr.12244.

Hanser, M. 2008. "The Metaphysics of Harm." *Philosophy and Phenomenological Research* 77 (2): 421–450. doi:10.1111/j.1933-1592.2008.00197.x.

Hanser, M. 2011. "Still More on the Metaphysics of Harm." *Philosophy and Phenomenological Research* 82 (2): 459–469. doi:10.1111/j.1933-1592.2010.00477.x.

Hanser, M. 2013. "Harm." In *The International Encyclopedia of Ethics*, ed. H. LaFollette, 2299–2307. Hoboken, New Jersey: Blackwell Publishing Ltd. doi:10.1002/9781444367072. wbiee186.

Holtug, N. 2002. "The Harm Principle." *Ethical Theory and Moral Practice* 5 (4): 357–389. doi:10.1023/A:1021328520077.

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

Lewis, D. 1979. "Counterfactual dependence and time's arrow." *Noûs (Detroit, Mich.)* 13 (4): 455–476. doi:10.2307/2215339.

Lewis, D. 2001. *Counterfactuals*. Malden: Blackwell Publishers.

McMahan, J. 1988. "Death and the Value of Life." *Ethics* 99 (1): 32–61. doi:10.1086/293034.

McMahan, J. 2002. *The Ethics of Killing. Problems at the Margins of Life*. Oxford: Oxford University Press. doi:10.1093/0195079981.001.0001.

Moore, G. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.

Norcross, A. 2005. "Harming in Context." *Philosophical Studies: An International Journal for Philosophy in the Analytic* 123 (1-2): 149–173. doi:10.1007/s11098-004-5220-3.

Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.

Scanlon, T. M. 1998. *What we owe to each other*. Harvard: Belknap Press.

Shiffrin, S. V. 1999. "Wrongful life, procreative responsibility, and the significance of harm." *Legal Theory* 5 (2): 117–148. doi:10.1017/S1352325299052015.

Shiffrin, S. V. 2012. "Harm and Its Moral Significance." *Legal Theory* 18 (3): 357–398. doi:10.1017/S1352325212000080.

Stalnaker, R. 1968. "A Theory of Conditionals." In: *Studies in Logical Theory*, edited by N. Rescher, 98–113. Oxford: Blackwell.

Thomson, J. J. 2011. "More On The Metaphysics of Harm." *Philosophy and Phenomenological Research* 82 (2): 436–458. doi:10.1111/j.1933-1592.2010.00418.x.