# Death does not harm the one who dies because there is no one to harm

An Epicurean riddle: What comes when we are not present, and when we are present it comes not?

## Introduction

The epigraph to this paper is, of course, derived from Epicurus when he says: ‘[D]eath, is nothing to us, since when we are, death is not present, and when death is present, then we are not’ (Epicurus in Taylor 2012a, p. 419). The premises of Epicurus’ argument are undeniable (everything after ‘since’), given what we might call the ‘termination thesis’, namely that death is the final end to our conscious existence (cf. Feldman 1994, Ch. 6). But there is a thought that he is equivocating on ‘nothing’ in the conclusion: ‘Death is nothing to us’. It is certainly true that in the sense of there being a subject to experience death there is no sensation of death—so in this sense it is certainly plausible to claim that ‘Death is nothing to us’. But in another sense, and the sense it is generally taken to mean, ‘nothing’ means that as there is no subject to experience death, death cannot *harm* one.

Much philosophical literature (ink spilt!) has been dedicated to showing how, exactly, death can be a harm to one who cannot experience any displeasure as a result. Philosophical imagination has run rife, concocting more and less plausible analogies to answer such questions as: how can one be harmed by something without experiencing anything unpleasant? (call this the *experiential* question) (see, e.g., Feldman 1994, Ch. 8; Fischer 1997; Nagel 1970; Taylor 2012b, Ch. 5); who, exactly, is it that is the subject of this harm? (call this the *existential* question) (see, e.g., Purves 2016); at what time does the harm of death afflict one? (call this the *timing* question) (see, e.g., Johansson 2016; Luper 2007; Taylor 2012b, pp. 81-3); what harm principle ought one to use? (see, e.g., Blatti 2012; Hanser 2008; Thomson 2011); and whichdeaths are bad? (see, e.g., Blatti 2012, pp. 320-3; McMahan 1988). My main concern in this paper will be to consider answers to the question on how one can be harmed by something without experiencing anything unpleasant, or the experiential question.

Those who claim that one can be harmed by something without experiencing anything unpleasant generally hold what is called a deprivation view (cf. Cyr 2017): it is the deprivation of any further intrinsic pleasures that is the cause of harm. The deprivation view is argued for in two main ways, either via a counterfactual claim that the deceased would have experienced more intrinsic good if they had not died or by providing a case that is analogous to being harmed in death. I will argue that both attempts fail. In particular, I first argue that Feldman’s attempt at providing counterfactual support for his claim, and Bradley’s strengthening of this with a Lewis-Stalnaker possible worlds account, fails by presupposing that their harm principle can be applied to the deceased. I then argue that the account of betrayal as analogous to death, put forward by Nagel and strengthened by Fischer, also fails. Specifically, I consider the debate between Fischer (1997, 2006, 2016) and Suits (2001, 2012), which has spanned three decades without any clear conclusion. I argue that part of the confusion, here, has to do with a conflation of a *de re* and a *de dicto* reading of betrayal and its associated harms. I argue that for *death* to be a harm it must be read *de re*. But the only plausible reading of betrayal is *de dicto*. So there is an equivocation in arguing that betrayals are analogous to death. Moreover, the only plausible reading of death’s harm—the *de re* reading—is not well justified, from the existent arguments. Thus, until better arguments are provided, one is best to agree with Epicurus and those who follow him that death is not a harm.

## The first strategy

Death strikes fear into the hearts of many. For some, it is the dying process that causes so much fear. For others, it is the state of affairs of one’s ceasing to exist. For others, it is the passing from the stage of dying to the state one is in after one ceases living that causes fear: will (the event of) death hurt? We might mean any one of these three things when we talk (think) about death. As Rosenbaum (1986) has pointed out, Epicurus is most charitably talking about the second: the state of affairs of one’s being dead. In any case, it is in this sense of death that the problem is so interesting; for, the thought goes, if one no longer exists to have sensations, what could it mean to say that one is harmed by one’s death?

Theorists have attempted to deal with this situation by pointing out that death is a kind of deprivation. Death, the thought goes, deprives one of what one could have had, namely life. In this sense, death is a harm not because of any positive features about death—for, what would they be?—but because of what death takes away, namely the opportunity to have more experiences (Nagel 1970). Whether this is a plausible account depends, as a necessary condition, on the possibility of one’s being harmed without experiencing anything unpleasant. Colloquially, the objection says that one cannot be hurt by something one does not know. Nagel states the objection as follows:

[D]oubt may be raised whether anything can be bad for a man without being positively unpleasant to him: specifically, it may be doubted that there are any evils which consist merely in the deprivation or absence of possible goods, and which do not depend on someone's *minding* that deprivation (Nagel 1970, pp. 75-6).

Philosophically, it rests on the hedonic premise:

1. An event or a state of affairs is a harm to a person only if it adversely affects her experiences (Taylor 2012b, p. 70)

(1) is at least plausible: it seems quite likely that it is a necessary condition for something to harm me that I be adversely affected by it, either now or in the future. This could come in the form of directly causing me some anguish or distress, or indirectly affecting my life as a whole. As my general well-being is determined by my experiences these can be called intrinsic goods\bads. There is nothing else in virtue of which they can be called good or bad.

The proponent of the deprivation view claims that there are some extrinsic bads which, while not necessarily intrinsically bad for the subject at the current time, lead to intrinsic bads in the future. For example, Effy’s turning down her job offer was extrinsically bad for her as it prevented any future promotions (we know this because she was marked: no future promotions), yet it was good for her at the time (she met her husband on her holiday). So while Effy’s turning down the job was not intrinsically bad (in fact, it resulted in an intrinsic good), it was extrinsically bad in virtue of its causing Effy’s life to contain less intrinsic value than it would have had if she had received future promotions. Other examples can be concocted, such as eating poisoned candy (Feldman 1994, pp. 133-4) or Jack’s putting rocks in Will’s cereal (Bradley 2009, p. 47). In these cases, these extrinsic bads are instrumentally bad in virtue of the intrinsic bad they will cause. In a principle:

CP: If something is extrinsically bad for a person, then it is bad for him or her because it leads to later intrinsic bads for him or her (Feldman 1994, p. 135).

While (CP) is plausible, it won’t get us all the way to the deprivation view. One complaint that may be made against (CP) is that it is impossible to tell, at the current time, whether something is extrinsically bad for a person. But this has to do with (CP)’s *predictive* power, and it never purported to be a predictive principle. It is enough for (CP) to be retrospective. All we can conclude is that there are some harms that are harms only in virtue of their effect on us in the future. The deprivation view holds that death is a harm because of what it deprives us of, so holds that death is an extrinsic bad (cf. Feldman 1994, pp. 135-6). But (CP) is inconsistent with this view, for there is nothing (according to (CP)) that it is bad *for*.

Feldman challenges (CP) on the grounds that it is an ‘overly narrow view’ in so far as ‘something must *cause* intrinsic evils for that person’ (Feldman 1994, p. 137). He proposes two examples to show that something may be overall bad for us even though it is not, and does not lead to any, bad experiences (p. 138). A representative example is as below:

Suppose a girl is born in a strange country—call it Country A. In Country A, they do not permit girls to learn to read and write. In this strange country, girls are taught to do laundry and raise children. Suppose this girl goes through life bearing children and washing laundry. Suppose she is reasonably satisfied, thinking that she has lived as a woman ought to live. She goes to her grave never realizing what she has missed. Suppose also that she had very considerable native talent for poetry—that she would have been a marvelously successful and happy poet if only she had been given the chance (Feldman 1994, pp. 137-8).

Feldman wants to conclude from this that while the woman does not experience any pain by being born in Country A, she has still suffered an intrinsic bad. Feldman (1994, p. 138) says that ‘it is a great pity that this woman had not been born in another country. I would say that something very bad happened to her, even though she never suffered any pain as a result’. If Feldman’s intuition is right, here, then it might show that (CP) is an ‘overly narrow view’ for ‘[s]ome things are bad for us even though they are not themselves painful experiences, and they do not lead to any painful experiences’.

Feldman uses this and similar examples to propose the following principle:

El: Something is extrinsically bad for a person if and only if he or she would have been intrinsically better off if it had not taken place (Feldman 1994, p. 138)

The example (of the girl) is intuitive and compelling but has assumed exactly what it was supposed to prove. It was supposed to answer the experiential question: can one be harmed by something without experiencing anything unpleasant? An answer to this question would explain what it *is* for one to be harmed without feeling pain. Examples of this abound: poisoned candy, breakfast rocks, etc. But rather than answering this (this-worldly) question directly, (EI) requires an answer to a preliminary question: what would a state where one is better off be like? So (EI) can explain how a deprivation can adversely affect one’s experiences *only if* this preliminary question can be answered. But to adequately answer this question requires one to reify unactualised possibilities; for it is just not clear how imagining the unactualised life I could have led, where I am intrinsically better off, makes it the case that I have been harmed in this actual life.

But perhaps I, being a middle-aged white heterosexual male, am not being sensitive enough to the example. The plight of under-privileged groups is surely a case for sensitive consideration, in respect to the bads they undergo due to lack of opportunity and oppression. It is not the imaginings of a better life, but the deficiency of the current one, which is relevant here.[[1]](#footnote-1) (EI), however, says little about the parameters needed to accurately use this principle. Let us, then, look at a more robust understanding of (EI), using Bradley’s Difference Making Principle:

DMP: The value of event E, for person S, at world *w*, relative to similarity relation R = the intrinsic value of *w* for S, minus the intrinsic value for S of the most R-similar world to *w* where E does not occur (Bradley 2009, p. 50).

(DMP) assumes a Lewis-Stalnaker possible worlds semantics. To determine which world is closest we try to keep as many things the same between both worlds as possible, e.g. a slight change in historical facts may be a closer world than one which has a change in the laws of nature. Consider the girl’s life. In one world, *A*, she lives in Country A, does laundry, raises children etc. Keeping as much as possible identical to *A*, there is another possible world, *B*, where *A* and *B* are identical (or near enough identical); however, in *B* the girl is born in Country B, where her creative arts can flourish. We need not envision a large change: a last-minute job offer for the father, or a long-term plan to relocate coming to fruition. It is from the point of this difference-maker that *A* and *B* diverge, her life being significantly different in each (but, of course, there will be some similarities). Let us agree with Feldman that *B* has more intrinsic value than *A* and so, all things considered, the girl would have been better off in *B*. Can we then draw the conclusion that the girl is harmed, given that *A* has less overall intrinsic value than its relatively similar alternative, *B*? The answer I will give is that (DMP) says ‘yes’ for the girl’s case, but this does not transfer over to the case of death.

On the Lewis-Stalnaker view the girl has a transworld identity. This is not a metaphysical claim—one cannot exist wholly in two different worlds—but a semantic one: the individual exists at, or according to, many different worlds (Divers 2007, p. 54; Lewis 1968, p. 114). Distinguishing between metaphysical and semantic commitment is helpful. We determine which possible worlds are relevant by using semantic content to give context. We are metaphysically (ontologically) committed to the relevant possible worlds and their constituents. This is what it is to be a realist about possible worlds. This would provide an answer to my concern about unactualised possibilities. These unactualised possibilities *are* actualised; they are actualised in the relevant possible world(s).

While this would successfully show how one can be harmed without experiencing any displeasure, its success is only due to combining (CP) and the counterfactual thought: ‘what would it be like to…’ Both of these claims are plausible, but the combination of them, when applied to death, is not. When buttressing counterfactual claims needed to make (IE) true, with (DMP), (CP)’s ‘narrowness’ can be seen to be due to its scope. In fact, (IE) is plausible *only* in virtue of (CP). However, when it comes to the deceased, (CP) is implausible. As (IE) gains its plausibility from (CP) and in so far as (DMP) is merely a fleshing out of (IE), we must (I will argue) reject this account of being harmed in death.

Using counterfactual reasoning seems to play some sort of role in planning for the future. For example, one might be motivated by the thought: ‘what would make me happier?’ and come up with a scenario, e.g. ‘my life would be improved if I read more Dickens’. We might then take measures to make this condition true (e.g. buying *Great Expectations*). We also use this kind of reasoning to learn from past mistakes, e.g. ‘my life would have been better if I hadn’t chosen to eat so much red meat’. The basic method is to imagine an improved state and determine the causal process needed to bring about this state. Sometimes improvement is impossible, especially when the improvement requires backward causation, choosing where one is born or whether or not one is deceased being prime examples. This is to be expected, however, given the mind’s resistance to (at least some) modal claims.[[2]](#footnote-2) How this works in practice is one of trial and error: sometimes we are wrong about what would make us happy, sometimes we provide ourselves with impossible targets.

(CP) draws a this-worldly conclusion from this-worldly facts. The problem encountered above was that it is not clear how this preliminary question—what would a state where one is better off be like?—could be answered. With the help of (DMP) this is now possible. Suppose *A* and *B* are as before. In *A*, the following counterfactual is (supposedly) true and made true by *B*: If the girl were born in Country B she would have been a marvelously successful and happy poet. The girl is harmed in *A*, therefore, because (we assume) her experiences were adversely affected. This is just an example of (CP): the girl’s birth in Country A is an extrinsic bad for her, due to her actual life being adversely affected. This derived an extrinsic bad—the girl’s birth in Country A, i.e. not in Country B—from an intrinsic bad—her actual life being adversely affected—via a counterfactual claim.

But consider what happens when we try to draw the analogy with death. This time, suppose @ is the actual world where *x* is dead and *w* is the closest possible alternative world where *x*’s counterpart—*x*C—lives another 10 (on the whole) good years. Feldman (1994, p. 140) tells us that ‘[death] is bad (when it is bad) because it deprives us of the intrinsic value we would have enjoyed if it had not taken place’. We can state this as the deprivation view:

Deprivation: *x*’s death (possibly) deprives *x* of many intrinsic goods[[3]](#footnote-3).

If (Deprivation) is true then it is made true by *x*C. *x*C experiences intrinsic goods not experienced by *x*. This is because of *x*C’s relative longevity. (IE) (and the strengthening of this in (DMP)) tells us that (therefore) *x* is harmed by her death. But (Deprivation) is a new move and requires support. Consider the closest possible world to @, *w*, where *x*Cis the person who most resembles *x*. *x*C makes *w* intrinsically more valuable for *x*C than @ was for *x*. So far so good. But nowhere has it been (successfully) argued that extrinsic bads can affect one completely independently of any intrinsic harm. (CP), recall, holds that φ is an extrinsic bad *only if* φ ultimately adversely affects the agent’s experiences (cf. (1)). But if φ is the agent’s death then *ex hypothesi* the agent has *no* experiences. So to conclude that death is an extrinsic bad because it prevents (or deprives) one of future intrinsic goods, where that one is *non*-*existent*, is a new claim.

Bradley attempts to support this claim with something like the following principle:

Prevention: Something could be extrinsically bad for a person without being intrinsically bad for her or him if it makes things worse than they would have been otherwise (cf. Bradley 2009, pp. 47-8).

Bradley (2009, pp. 47-8) provides the following thought experiments to support his view:

If Jack receives a vaccination that prevents him from contracting a painful illness, the vaccination is instrumentally good for Jack, even if it does not cause him any enjoyment. Likewise, if Jack intercepts some money intended for Will without Will ever finding out about it, and the money would have made Will happy, Jack's action is instrumentally bad for Will, even though nothing intrinsically bad happens to Will as a result of the action.

Both these instances can be explained with reference to (CP) and the counterfactual strengthening of this claim. Jack’s vaccination is instrumentally good for Jack in virtue of *Jack’s good health*, comparative to the intrinsic value of the world where, if Jack were not to have the vaccination, he would have gotten sick. Jack’s interception is instrumentally bad for Will because it prevents Will from having the money he would have had if Jack had not intercepted. If (CP) and our counterfactual reasoning can give us this conclusion we don’t need (Prevention) to explain these intuitive results.

Moreover, (Prevention) implicitly relies on one’s being adversely affected for the attribution of a value-claim. We would usually say that it is extrinsically bad for me to forget to set my alarm (or, if you are worried about a negative fact causing something: set my alarm an hour later) in order to prevent me from sleeping in and missing some (let us say, insignificant) appointment. So not setting my alarm is an extrinsic bad. Now, suppose there was nothing in virtue of which I did not set my alarm. Am I still to count this as an extrinsic bad? Bradley would say ‘no’, for while it satisfies the condition that something might be extrinsically bad without being intrinsically bad, it does not satisfy the condition that it makes things worse than they would have been otherwise. But what can this condition mean except for adversely affecting my *experiences* (comparatively); for there is nothing else in virtue of which we can distinguish between the extrinsic bad of my setting the alarm late when I need to arise at a certain time and the value-less instance of me not setting my alarm. Hence, the difference between these two examples depends on *how I am affected*.

To deny that the value of one’s life does not ultimately rest on (CP), and how one’s experiences are affected, one must deny the hedonic premise, i.e. (1). But doing so means one has changed the subject and is using a *different* understanding of intrinsic value, such as desire or preference satisfaction. A prime example of such a case is this.[[4]](#footnote-4) Suppose Jed is a scientist who predicts that there used to be life on Mars. In fact, he dedicates his life to proving this conjecture. Unfortunately for Jed, just before the mission departed Earth he died of heart problems. The mission went exactly as planned. At this point, the intuition may be that if the mission was successful Jed’s life would have been significantly more valuable (after all, this is what his entire life was dedicated to). If it brought back conclusively negative results, one might think, Jed’s life would have been significantly *less* valuable. But this intuition just denies (1): *Jed* is unaffected by the results, while his desires or preferences (while he was alive) may not be.

One may object by saying that as I am committed to (DMP) and (DMP) implies (Prevention), I must accept (Prevention). Worse, because I say that (CP) implies (DMP), if I were to claim (Prevention) is false I must, via Modus Tollens, say that (CP) is false. But this is wrong on both counts. First, this gets the argument structure wrong. As I noted, I have some concerns about using unactualized possibilities (counterfactual claims) to make this-worldly conclusions, so I reject the claim that (CP) implies (DMP) (or (IE), for that matter). Rather, if (*if*) (DMP) is true then it requires (CP) as a necessary condition. There may be a problem, however, if (DMP) implies (Prevention) because (DMP) would imply two contradictory principles. But (DMP) does not imply (Prevention). Why is it more plausible to say that *x*C is the counterpart of *x*,after the latter’s death, rather than someone else, very similar to *x*? Perhaps *x*C is very similar to *both x and y* (both of who are in @). Suppose *x*’s death occurs at *t*. Why is it not the case that prior to *t*, *x*C is a counterpart of both *x* and *y*,whereas after *t*, *x*C is a counterpart of *only y*. But why include *y* at all? Why is it more plausible that *x* is a counterpart of *x*C after *t* rather than *x*C having no counterpart in @ at all? The decision that *x*C is a counterpart of *x* after *t* is unsupported. Given the termination thesis—that death is the final end to our conscious existence—there is stronger support for the claim that *x*Chas *no* counterpart in @.

The experiential question still remains. The first attempt at arguing for the deprivation view via a counterfactual claim fails due to its presupposing that one can be harmed while experiencing no displeasure as a result.

## The second strategy

Nagel reasons that if the claim that one cannot be harmed without experiencing any displeasure is a valid objection to death being a harm, then it will apply to many other harms. He then argues that there are some other harms it does not apply to. A prime example of this is betrayal:

The first type of objection is expressed in general form by the common remark that what you don't know can't hurt you. It means that even if a man is betrayed by his friends, ridiculed behind his back, and despised by people who treat him politely to his face, none of it can be counted as a misfortune for him so long as he does not suffer as a result… It seems to me worth asking what assumptions about good and evil lead to these drastic restrictions. (Nagel 1970, p. 76)

But, as has been pointed out by various philosophers, this example may fall short of showing that *death* could be a misfortune (Fischer 1997; Rosenbaum 1986). The case of betrayal may, in fact, *not* be analogous to death; for while the betrayed (according to Nagel) does not suffer as a result of his misfortune it is surely possible he *could*. In death, however, the deceased is *incapable* of experiencing the misfortune (assuming it is a misfortune). For the deceased but not for the betrayed, the counterfactual—could find out about the misfortune—is *always* false.

In light of these objections, Fischer has proposed a revamped betrayal case. Fischer proposes a Frankfurt-style version of Nagel’s betrayal case:

Imagine first that the example is as described by Nagel. You are betrayed behind your back by people who you thought were good friends, and you never actually find out about this or have any bad experiences as a result of the betrayal. But now suppose that these friends were (very) worried that you might find out about the betrayal. In order to guard against this possibility, they arrange for White to watch over you. His task is to prevent you ever from finding out about the betrayal (Fischer 1997, p. 345).

How White achieves this need not waylay us. It is enough to suppose that every time one of your good friends is about to let something slip White prevents this. Every time one of your friends is about to act in such a way that you are disadvantaged by their betrayal White steps in. Perhaps he does this by diverting your friend’s attention at the crucial point. Being a Frankfurt-style case, however, means that White need not step in at all. It is enough that you do not ever find out about this betrayal. White is merely a pre-caution that *ensures* you do not find out about the betrayal. The White betrayal case merely highlights the point that betrayal harms the one betrayed even if the counterfactual ‘could find out about the betrayal’ is false. If the counterfactual ‘could find out about the betrayal’ is false you are incapable of finding out about this betrayal. As such, the White case appears analogous to death. The intuition behind this case is that one is harmed (by betrayal) even should one be incapable of experiencing any displeasure as a result.

There are two (not mutually exclusive) ways this analogy might fail. It may be an inappropriate analogy of death or, while being an appropriate analogy, it may not work in the way Fischer thinks it does. Taylor (2012b, Ch. 5) has argued that the accounts put forward by Nagel, Feldman and Bradley all fail because they do not provide appropriate analogies of death. But while Taylor’s objection may work for the strategy adumbrated above, it will not work for the current one. This provides support for considering this second strategy as independent of the first.

Taylor reasons that, in the analogy, the subject either exists or does not exist: if the former, then the analogy is fundamentally different to death; if the latter, then the question is begged against the Epicurean (Taylor 2012b, p. 78). But this move is too quick. The first strategy I considered takes Taylor’s second lemma, so I need not consider that here; but Taylor is wrong to point out the disanalogy when this is applied to *betrayal*. In death one has a zero-level experience for the simple fact that one is incapable of experiencing *anything*, whereas in the betrayal case one has a non-zero-level experience: one still experiences soft summer breezes, etc. But while there is a disanalogy between the existent subject in the betrayal case and the non-existent subject in the death case, Taylor’s first lemma overlooks the pertinent analogy, which is the zero-level experience of the *betrayal*. The deceased has a zero-level experience of everything, which includes zero-level experiences of particular events, such as soft summer breezes and betrayals. The betrayed has zero-level experiences of only particular events, such as those events that are impossible, e.g. being in two places at once (logical impossibility), or running a mile in under a minute (physical impossibility). The White style betrayal case is one where the betrayed has a zero-level experience of a particular event. As the deceased has zero-level experiences of particular events (namely all of them) and the non-deceased has zero-level experiences of certain events—the White betrayal case being a pertinent example—the analogy seems appropriate.

Suits has responded to Fischer on betrayal, and similar cases, on numerous occasions (see Suits 2001, 2012). He objects to the idea that betrayal *as such* is a bad thing for an individual, as the action we call a ‘betrayal’ might result in good consequences or no consequences at all. For example, suppose a friend ‘betrayed’ you by spreading rumours about your being able to hide your deceit, which resulted in your being hired for your dream job: a political post, perhaps.[[5]](#footnote-5) In the White-monitored betrayal case it is stipulated that there are no bad consequences. As these examples show, betrayals do not necessarily lead to bad consequences. The conclusion Suits wishes to draw is that the reason we say that actions we are apt to call betrayals, and similar cases, harm the individual are the *consequences* of the act, i.e. the experiencing of displeasure, rather than the nature of the actual act, itself (see Suits 2012, p. 225). This would be a win for the Epicurean.

Fischer denies this reasoning and claims that it is the nature of betrayal, itself, that shows a betrayal is harmful (see, e.g., Fischer 2006, p. 363). Fischer seems to be following Nagel, here. Nagel’s opponent is ‘[s]omeone who holds that all goods and evils must be temporally assignable states of the person’, where things such as ‘[l]oss, betrayal, deception, and ridicule are… bad because people suffer when they learn of them’ (Nagel 1970, p. 76). So Nagel’s target seems to be someone like Suits, who thinks that betrayals are bad only in virtue of their consequences. (Indeed, this is what the hedonist ought to say.) Nagel thinks we ought to ‘accommodate these cases directly’ by asking ‘how our ideas of human value would have to be constituted’. In a later version of his 1970 paper he writes: ‘the natural view is that the discovery of betrayal makes us unhappy because it is bad to be betrayed – not that betrayal is bad because its discovery makes us unhappy’ (Nagel 1979, p. 5).

This way of arguing the point will end in a stalemate (and, indeed, it has if the debate between Suits and Fischer is a representative example). And here lies the catch: in effect, Nagel is denying the hedonic view of value (in (1)). For example, Nagel might agree that death is not a harm on the hedonist view, but claim that the hedonist view is the wrong way to think about human value. In this way he would be using our intuitions about the harm of death to argue for his view of human value. But if the analogy is used in this way then it will not succeed in its goal of showing that there is an analogy of death, for it merely assumes what it was supposed to show: we are reconciling our theoretical framework with our intuitions, rather than believing based on logic and rational argument.

The way to reconcile this issue is to reconceptualise it as one, not between different conceptions of human value (for this ends in a stalemate), but as one between different conceptions of betrayal. To do this, let us quantify what is meant by harm and distinguish between a *de re* and a *de dicto* reading of ‘betrayal’. Consider Suits’ (2012, p. 226) conditions of when one suffers some harm (on the hedonist view):

1. suffering,
2. the risk of suffering,
3. being denied better experiences
4. the risk of being denied better experiences.

Call these the harm conditions, or H, where H is the disjunction of (i)-(iv). With the use of these harm conditions, ‘betrayal’ can be given a *de re* or *de dicto* reading. To do this, it is helpful to think of the modal properties of ‘betrayal’ in terms of possible worlds. If betrayal generally harms the individual then it had better be the case that it generally harms the individual in every world in which a counterpart of the betrayal takes place. Taking on board lessons learnt from Jon Tresan (2006), Fischer’s claim that ‘betrayal’ will always be a harm to the betrayed can be given a *de re* or *de dicto* reading, depending on where the necessity operator is placed. To say that betrayals generally harm the one betrayed is to say, either:

B(*de* re): If *x* betrays *y*, then, necessarily, x’s betrayal of y generally results in H

B(*de dicto*): Necessarily, if *x* betrays *y*, x’s betrayal of y generally results in H

A *de re* reading is concerned with the actual nature of the thing in question; a *de dicto* reading is concerned with the concept under which the thing can be classified. There will be instances where not both of these readings are true. To clarify matters, consider a different claim about a particular concept. Consider a *de re* and *de dicto* claim about research assistants:

R(*de* re): Research assistants necessarily do research

R(*de dicto*): Necessarily, research assistants do research

Here, R(*de dicto*) is a claim about the concept ‘research assistant’. It provides an explanation of the concept ‘research assistant’ and is true: if *x* is a research assistant then *x* will do research. R(*de re*), on the other hand, is a stronger claim. It tells us that the particular individual who happens to be a research assistant couldn’t but do research. But this is false. Suppose Fred is a research assistant in @ and he does research. So R(*de dicto*) is true of Fred. R(*de re*), however, is not; for there is a possible world, *w*1, where a counterpart of Fred exists but he is a Cook. So it is not true that if *x* is a research assistant then *x* will necessarily do research. Whereas R(*de dicto*) is a claim about our concepts, R(*de re*) is a claim about particular individuals. R(*de dicto*) could be made false if we use the concept ‘research assistant’ to apply to cases where research assistants do not do research.

Suits’ argument is most plausibly directed at B(*de dicto*). This can be best seen when Suits compares the case of betrayal with another risky behaviour, namely incautious firings of guns (Suits 2001). In normal circumstances, i.e. when there are no safeguards in place, we would say that an incautious firing of a gun is risky due to the harm it might cause. But if the appropriate safety precautions are taken, to mitigate any risk, then incautious firings of guns are no longer possible causers of harm, and so are no longer risky. Here, it is the *consequences* of the action or event that determines whether or not it is risky. He thinks that the same is true of betrayals, i.e. it is the consequences of the action that makes a betrayal bad, rather than betrayal, itself, being bad. This is the converse of what Nagel said (above) on betrayals.

In his 2001, Suits puts the point in the following way:

Incautious firing of guns is risky, but once the precautions are in place, then firing the gun is not at all incautious. Similarly, betrayals are risky, but once White is in place then there is no betrayal after all. (Perhaps there was an attempt at a betrayal; and we may thank the good Mr. White for standing by to prevent any possible harmful consequences as a result of the attempt.) (p. 77).

Here, Suits identifies an action under a description, i.e. betrayals and incautious firings of guns, and attributes the concept of *risky* to actions that fall under these descriptions. He reasons that if an action-under-a-description is not risky, the action cannot thereby fall under the description of an action that is risky, e.g. ‘betrayal’ or ‘incautious firing of guns’. This is what one would expect if Suits were concerned with a *de dicto* reading of ‘betrayal’. B(*de dicto*) relates to how we use the concept ‘betrayal’: we call an action a betrayal when the action generally results in H. That this is the best way to understand Suits can be corroborated when he says that ‘once White is in place then there is no betrayal after all’ (Suits 2001, p. 77). This is because ‘betrayal’ is understood as denoting an action where H generally ensues, but if there is no chance of H ensuing we are apt to withdraw our attribution of ‘betrayal’. This is similar to how Fred’s notgenerally doing research means we would not call him a research assistant.

Fischer’s conception of betrayal, on the other hand, appears to be *de re*. In his (2006), Fischer concedes to Suits that incautious firings of guns could, indeed, not be risky if appropriate measures are in place. But he insists that betrayals are *always* harmful. He justifies this by claiming that the ‘contention is not that all riskless behavior harms others (or that all riskless behavior which would be risky, but for the precautions[] harms others[)]; the contention is only that some does’ (Fischer 2006, p. 363). So here he agrees with Nagel, that we call betrayals bad because they are betrayals and not for their consequences. But to justify this move Fischer needs to provide something other than a different conception of human value. The account Fischer gives, however, is insufficient. The closest Fischer comes to providing a justification for this distinction is as follows:

[T]he cases are importantly different. Whereas I am inclined to say that the negative characterizations of you by your so-called ‘‘friends’’ harm you, *by the very nature of the behavior* [my emphasis], I have no similar inclination to say that a mere firing of a gun, where there is no chance of hitting anyone, can harm anyone (except, perhaps, the gun-firer, but that raises different issues...) (Fischer 2006, p. 363).

Here, it is ‘the very nature of the behaviour’ that distinguishes betrayals from incautious firings of guns. Talking about the *nature* of the behaviour, rather than the *type* of behaviour, is evidence of a *de re* interpretation, as the speaker is concerned with the very nature of the action undertaken rather than the concept under which this action falls. This is similar to how calling Fred a research assistant need not be revoked when Fred assumes other duties, if it is understood that the attribution ‘research assistant’ denotes Fred’s *nature* (whatever this might mean).

But, to echo the words of Suits (2012, p. 225), Fischer has merely given a name to this behaviour and has not explained why it differs from incautious firings. Surely the point can be granted that there is something about the very nature of betrayals that make them betrayals, and which distinguish them from incautious firings; but this is only to distinguish incautious firings from betrayals and says nothing about the inherent riskiness of each behaviour.

Fischer offers a response to this worry, in his (2016). This response appears to explain the *nature* of betrayals, where this has to do with a ‘normative theory of human interests’ (p. 141). He thinks that we have an interest in not having our reputations destroyed, even if we do not find out about this. He does not think the same is true for incautious firings of guns:

I have the intuition that a White-monitored verbal assault does indeed harm the target of the assault. But I do not have the intuition that an incautious firing of a gun, in a context in which it is impossible for (say) a particular individual to be affected by the gun firing, is a harm to the individual in question. The incautious firing of a gun is wrongful behavior – it is either reckless or negligent. But it does not follow that it harms anyone (Fischer 2016, p. 141).

Fischer continues:

I’m inclined to say that the proper response might rely on a normative theory of human interests. It is plausible that we have an interest in not having our reputations besmirched, even if we never find out about these besmirchings… Perhaps, then, the way to explain the asymmetry in our intuitive reactions to riskless incautious gun-firings… on the one hand, and riskless betrayals, on the other, is in terms of the implications of the most plausible normative theory. Such a theory would posit an asymmetry in the interests of human beings (pp. 141-142).

What might these interests be? Presumably they will differ depending on the behaviour. Thus, for Fischer, this ‘asymmetry’ will consist in our caring more about our reputations being besmirched than incautious firings of guns. But does this explain the nature of betrayals? It is not clear that it does. *Some* betrayals will consist in a besmirching of our reputations but not all will. Consider the case of infidelity: this seems to be an instance of betrayal yet is not a besmirching of one’s reputation. Rather, it is an example of a discordance between how the world is and how the betrayed perceives the world. The harm consists, therefore, in the betrayed’s personally important beliefs being rendered false (cf. Hetherington 2001, p. 351). This is particularly harmful as it may make false the agent’s legitimate, normative-changing expectations of how the world is or ought to be. This is merely an instance of the ultimate betrayal (if it is a betrayal), namely the belief in God.

There are thus two different conceptions of the nature of betrayals. Plausibly, each might be subsumed under the other. It might be that a discordance between how the world is and our beliefs about the world vie with our interests, or it might be that every besmirching can be understood as a discordance between our beliefs about our reputation and how others see us. Fischer prefers his account to Hetherington’s because if the badness of betrayals arise due to a discordance, then the analogy with death fails due to the non-existence of the subject, in the latter case. But once this move has been made Fischer has lost the debate. For one to have an interest in φ requires one to exist, similarly to how a discordance requires a subject. To establish this view, therefore, requires Fischer to support a direct account of death’s evil (rather than merely showing an analogue). This is because, intuitively, if ψ’s badness depends on the thwarting of one’s interests then there must exist a subject whose interests can be thwarted. Specifying that ψ is death does not change this conclusion without argument. As such, this view’s success is dependent on the first strategy I considered. Given the failure of the first strategy we must thus conclude that this second strategy fails, also.

If a good way to settle an argument is to say that you’re both right, a bad way to settle an argument may be to say that you’re both wrong. So let’s settle this argument badly. If Suits has a *de dicto* reading of ‘betrayal’ then there is a response Fischer can give. But Fischer is right to concentrate on B(*de re*) because without the *de re* reading of ‘betrayal’ the analogy with death does not work. However, (I will argue) B(*de re*) is implausible and thus cannot be used as a way to answer the experiential question in the affirmative: can one be harmed by something without experiencing anything unpleasant? In short, Suits’ complaint can be answered and Fischer’s solution is implausible.

B(*de re*) is implausible. For B(*de re*) to be true it must be true, for any action that is a betrayal in @, that that action generally result in harm in all accessible possible worlds. But this is necessarily false in White-monitored worlds. There are other examples, too, where betrayals result in good consequences rather than bad, in close possible, if not the actual, world(s), e.g. one’s receiving a political post as a result of the betrayal. The problem, here, is general: we conflate the value of an action with the expected consequences of that action, where the action has no value at all. Why would anyone think that betrayals are harmful? It is not *that* the action is a betrayal which makes it harmful, but the nature of particular actions, themselves, from which we associate betrayals with harm. Concepts are concepts and actions are actions. It is not concepts that harm or scare us, but what the action (event or whatever) the concept denotes that strikes fear in our hearts. This feeling of fear from the action can leave traces on the word, itself. Hence, a high school student’s surprise when they encounter their University teachers and find that ‘teacher’ is a term with a denotation and not a connotation. This is a kind of ‘word magic’, often utilised by advertising and media companies to draw in consumers. It has real causal bif (depending on one’s social milieu) but this ‘bif’ is dependent on its connotations and is socially entrenched. This seems to point to the conclusion that whatever harm is associated with ‘betrayal’ is due to the connotation of this concept, rather than an intrinsic feature of betrayals, *as such*. To conflate the connotations of the word with its intrinsic features, in this instance, is to resort once again to values, similarly to the way Nagel did.

While B(*de re*) is implausible, there is a way for Fischer to justify B(*de dicto*), *contra* Suits. It could be the case that if φ is a betrayal it is always generally harmful. There is a concern, however, that this reading of betrayal may be false for White-monitored verbal assaults. Following Suits (in his 2001), this might be a case where the concept ‘betrayal’ does not apply. However, White-monitored cases are anomalies. Couldn’t a no Frankfurt-style-counterfactuals clause be included in B(*de dicto*)?:

B(*de dicto*)no Frankfurt: Necessarily, if *x* betrays *y*, x’s betrayal of y generally results in H (assuming there are no Frankfurt-style-counterfactuals)

The no-Frankfurt-style-counterfactuals amendment is not a cheat. Compare the original Frankfurt case. For one to have free will the following condition must be satisfied ‘could have done otherwise’. The Frankfurt case shows that our intuitions say that this condition might not be satisfied yet we still act freely. Similarly, Fischer wants to say that one could be betrayed even if the following condition does not hold: ‘could find out about the betrayal’. Surely common betrayals are not generally White-monitored; hence, it seems that a harm will generally result, even if it does not do so in this actual instance. Fischer thus wants us to draw the conclusion, and he seems right in this, that betrayals are harmful even if one could not find out about the betrayal.

This works because B(*de dicto*) has to do with how one would generally apply the concept to a certain sort of behaviour, namely betrayal-type behaviour. That no one actually experiences any displeasure as a result (cf. White-monitored verbal assaults) is irrelevant to the *type* of behaviour it is. In other words, it turns out to be true that in all possible worlds where one is betrayed one is generally harmed. However, it is not true, in all possible worlds, that the action we associate with betrayal-type behaviour harms the individual: there are (possible, if not actual) instances where the betrayed is never impacted by the betrayal. This means that ‘you are betrayed by your friends’ and ‘you could not be harmed by your friends’ actions’, could be true simultaneously, because the truth of these sentences depend on different readings of ‘betrayal’: the former depends on B(*de dicto*) while the latter depends on B(*de re*) (or, if not, B(*de dicto*)no Frankfurt).

This move can solve the discrepancy in Fischer’s intuitions (on the betrayal\riskless incautious firings case) without having to refer to a theory of normative interests. To do this, he simply holds B(*de dicto*)no Frankfurt and a *de dicto* reading of ‘incautious firings’. Is this a cheat? No and yes. It is not a cheat if he is comparing White-monitored betrayals and incautious firings, as B(*de dicto*)no Frankfurt is not a cheat. But it is a cheat in the sense that the comparison is not equivalent. An equivalent comparison would require an explicit harm-free betrayal. But, for Fischer, this would be a contradiction in terms, if he subscribes to B(*de dicto*)no Frankfurt, which I have argued he needs to. So it seems that Fischer is merely led by his intuitions that betrayals are intrinsically harmful.

## In summary

Rather than having two different intuitions about whether betrayals are examples of one’s being harmed without feeling any displeasure as a result, we have intuitions about two different readings of betrayal. As long as ‘betrayal’ continues to mean betrayal, as determined by its use in the English language, Fischer’s example of White seems sufficient to show that one can be harmed without one’s coming to know of the betrayal. So this much seems to succeed. But this is a conclusion about our *concepts* and not about the actual nature of the behaviour (as Fischer claims it to be). It does show that one can be harmed in a betrayal. And it does show that there is a token of the type ‘one can be harmed by something one cannot experience as being unpleasant’. (This is what the example was originally supposed to show.) But it does *not* show that there is some action that is intrinsically harmful, due to its very nature. One would need B(*de re*) for this claim.

B(*de dicto*) is too weak to draw the analogy with death. For the analogy to work one requires B(*de re*), but B(*de re*) is implausible and rests on a kind of ‘word magic’. B(*de dicto*) has to do with the way we use ‘betrayal’ rather thanthe actions (or events) that fall under this concept. However, what is important to me is whether my actual death will harm *me* and not whether the *concept* of death implies that the deceased is harmed. The proposition that needs to be shown true is: Death is harmful to the one who is dead. There are two readings one might have of this claim:

D(*de re*): There is someone of whom death harms (or, if *x* dies then, necessarily, *x* is generally harmed)

D(*de dicto*): Death generally results in harming the individual (or, necessarily, if *x* dies then *x* is generally harmed)

The pertinent concern I have is whether or not D(*de re*) is true. How I use the terms I use does not strike fear into my heart. But all those who use an analogy similar to the betrayal case can do is show that D(*de dicto*) may be true. But this relates to how we use our concepts.

An extrapolation of Fischer’s (and Nagel’s) argument from betrayal shows the problem (cf. Fischer 2016, p. 146; Nagel 1970, p. 76):

1. If the following is true: If *x* dies then, necessarily, *x* is generally harmed, then one can be harmed without feeling any displeasure as a result (*de re* reading of death’s harm)

2. This is implausible unless there are other instances of one’s being harmed without feeling any displeasure as a result

3. ‘Betrayal’ is an instance of one’s being harmed without feeling any displeasure as a result (*de dicto* reading of betrayal’s harm)

Therefore, it is not implausible that death is a harm\*

As can be seen with this reconstruction, whether one can get the conclusion depends on Fischer’s move in (3). To get this conclusion one needs a *de re* reading of ‘betrayal’;(3), however, relies on a *de dicto* reading and thus the conclusion does not follow. As this strategy to argue for the deprivation view of death can only suggest that death may be intrinsically harmful by analogy with other intrinsic harms, and the prime analogy equivocates between a *de re* and a *de dicto* reading of the concept and its respective harms, the experiential question, namely, how one can be harmed by something without experiencing anything unpleasant, remains unanswered.

## Conclusions drawn from this discussion

Drawing these conclusions together, it can be seen that for death to be intrinsically harmful requires a *de re* reading of death’s harm (what have I to fear in a concept?). A direct route to this conclusion was proposed via a counterfactual claim that the deceased would have experienced more intrinsic good if they had not died. It was shown, however, that Feldman’s argument for this conclusion, and Bradley’s strengthening of this in (DMP), failed due to the non-existence of the subject. The second strategy avoids this problem by providing a case that is analogous to being harmed in death. It was seen, however, that the argument for this conclusion equivocates on harm in betrayal (B(*de dicto*))and harm in death (D(*de re*)).

Given that we determine how we use our concepts, it is up to us how we use the term ‘death’. That is, we are pressed with a different question than the one we started with: ‘*Ought* we use “death” in a way that is necessarily (*de dicto*) related to harm?’ Thus, we are led straight back to the anxiety Epicurus was attempting to calm: ought we be fearful of death? The answer to this question cannot be ascertained. But surely the question is urgent, cannot be ignored. Given the compulsion to answer this question one way or another, therefore, it may come down to a pragmatic response. Hence, given that it is better to live without fear than with, it is better to conclude that death is *not* a harm. And why would one say any different?

## References

Blatti, S 2012, 'Death's distinctive harm', *American Philosophical Quarterly*, vol. 49, no. 4, pp. 317-30.

Bradley, B 2009, *Well-Being and Death*, Oxford University Press, Oxford Scholarship Online.

Cyr, TW 2017, 'How does death harm the deceased?', in JK Davis (ed.), *Ethics at the end of life: new issues and arguments*, Routledge, New York, pp. 29-46.

Divers, J 2007, 'Quinean scepticism about de re modality', *European Journal of Philosophy*, vol. 15, no. 1, pp. 40-62.

Feldman, F 1994, *Confrontations with the reaper: a philosophical study of the nature and value of death*, Oxford University Press, New York.

Fischer, JM 1997, 'Death, badness, and the impossibility of experience', *The Journal of Ethics*, vol. 1, pp. 341-53.

—— 2006, 'Epicureanism about death and immortality', *The Journal of Ethics*, no. 10, pp. 355-81.

—— 2016, 'Mortal harm', in S Luper (ed.), *The Cambridge companion to life and death*, Cambridge University Press, Cambridge.

Hanser, M 2008, 'The metaphysics of harm', *Philosophy and Phenomenological Research*, vol. LXXVII, no. 2, pp. 421-50.

Hetherington, S 2001, 'Deathly harm', *American Philosophical Quarterly*, vol. 38, no. 4, pp. 349-62.

Johansson, J 2016, 'When do we incur mortal harm?', in S Luper (ed.), *The Cambridge companion to life and death*, Cambridge University Press, Cambridge, pp. 149-64.

Lewis, D 1968, 'Counterpart theory and quantified modal logic', *Journal of Philosophy*, vol. 65, no. 5, pp. 113-26.

Luper, S 2007, 'Mortal harm', *The Philosophical Quarterly*, vol. 57, no. 227, pp. 239-51.

McMahan, J 1988, 'Death and the value of life', *Ethics*, vol. 99, no. 1, pp. 32-61.

Nagel, T 1970, 'Death', *Nous*, vol. 4, no. 1, pp. 73-80.

—— 1979, 'Death', in T Nagel (ed.), *Mortal questions*, Cambridge University Press, New York, pp. 1-10.

Purves, D 2016, 'Accounting for the harm of death', *Pacific Philosophical Quarterly*, vol. 97, pp. 89-112.

Rosenbaum, SE 1986, 'How to be dead and not care: a defense of Epicurus', *American Philosophical Quarterly*, vol. 23, no. 2, pp. 217-25.

Smutts, A 2012, 'Less good but not bad: in defense of Epicureanism about death', *Pacific Philosophical Quarterly*, vol. 93, pp. 197-227.

Suits, D 2001, 'Why death is not bad for the one who died', *American Philosophical Quarterly*, vol. 38, no. 1, pp. 69-84.

—— 2012, 'Death and other nothings', *The Philosophical Forum*, vol. 43, no. 2, pp. 215-30.

Taylor, JS 2012a, 'Bioethics and the metaphysics of death', *Journal of Medicine and Philosophy*, vol. 37, pp. 417-24.

—— 2012b, *Death, posthumous harm and bioethics*, Routledge, New York.

Thomson, JJ 2011, 'More on the metaphysics of harm', *Philosophy and Phenomenological Research*, vol. LXXXII, no. 2, pp. 436-58.

Tresan, J 2006, 'De Dicto Internalist Cognitivism', *Nous*, vol. 40, no. 1, pp. 143-65.

Warren, J 2004, *Facing Death*, Oxford University Press, Oxford Scholarship Online.

1. There are two relevant points to make here, one concerning the subject’s possible goods and the other concerning the relevant comparison. As to the first, Blatti (2012, pp. 321-2) points out (following McMahan 1988, p. 41) that ‘the badness of death is determined not by the deprivation of goods the victim *might* have enjoyed, but by whatever goods the victim in fact would have enjoyed’. Smutts (2012, pp. 205-11; cf. Warren 2004, pp. 33-4) points out that the comparative account would give us the wrong result if it conflates less good states, e.g. winning $100,000 instead of $1,000,000, with bad ones, e.g. losing one’s life savings; for it would (un-intuitively) conclude that winning $100,000 would be intrinsically bad for the subject. [↑](#footnote-ref-1)
2. This is a general claim and seems plausible. As such, I want to avoid the issue of whether, e.g., the mind’s clearly conceiving *x* means that *x* is possible, which seems to involve more than just idle wistful thinking, like that suggested here. [↑](#footnote-ref-2)
3. The modal operator is required here, *contra* Nagel (1970, pp. 79-80), due to some considerations about the rationality of suicide and euthanasia, i.e. cutting one’s full-of-suffering life short (McMahan 1988). [↑](#footnote-ref-3)
4. A similar example was proposed to me by Patrick Stokes. [↑](#footnote-ref-4)
5. For a different example of one benefiting from a betrayal-type action, see Suits (2012, p. 225). [↑](#footnote-ref-5)