Not All Killings Are Equally Wrong

Todd Karhu

Ronald Dworkin once observed that "most people think . . . that the murder of a depressive handicapped octogenarian misanthrope is as heinous, and must be punished as seriously, as the murder of anyone younger or healthier or more valuable to others."[1] Jeff McMahan, a prominent defender of this view, calls it the *Equal Wrongness Thesis*.[2] As he states the thesis, it holds that the strength of the moral objection to killing persons "does not vary with such factors as the degree of harm caused to the victim, the age, intelligence, temperament, or social circumstances of the victim, whether the victim is well liked or generally despised, and so on."[3]

Whether the Equal Wrongness Thesis is true is of obvious theoretical interest. It also has important practical implications, since any factor that is irrelevant to the strength of the moral objection to an act of killing is *ipso facto* irrelevant to that act's overall permissibility. Suppose we must decide, as the German Federal Constitutional Court once did, whether to pass legislation permitting the gunning down of hijacked passenger planes when doing so is necessary to prevent their use as a terrorist weapon.[4] If the strength of the moral objection to killing someone is unaffected by how bad it would be for her to die, then it makes no difference to the permissibility of shooting down these aircraft that the passengers on board them would almost certainly have died soon anyway.

My aim in this article is to argue against the Equal Wrongness Thesis by appeal to a moral equivalence between killing someone

---

[1] Dworkin 1993, p. 85.

[2] McMahan 2002, p. 235. See also McMahan 1995.

[3] McMahan 2002, p. 235.

[4] Bundesverfassungsgericht, *Judgment of the First Senate of 15 February 2006*. The court ruled against the proposed legislation, arguing that "[h]uman life and human dignity enjoy the same constitutional protection regardless of the duration of the physical existence of the individual human being" (paragraph 132).

and knocking someone unconscious. The argument has two premises, each of which asserts the moral equivalence of two types of act. The first premise holds that depriving someone of a certain period of conscious life by killing her is as wrong as depriving someone of the same amount of conscious life by knocking her unconscious until her death. The second premise holds that depriving someone of a certain period of conscious life by knocking her unconscious until her death is just as wrong as depriving someone of the same amount of consciousness by knocking her unconscious in the middle of her life.

The truth of these two premises is enough to undermine the Equal Wrongness Thesis. From the first premise, we can have it that killing a person a month before she would have died is as wrong as knocking someone unconscious for the last month of her life. And from the second premise, we can have it that knocking someone unconscious for the last month of her life is as wrong as knocking someone unconscious for a month in the middle of her life. The Equal Wrongness Thesis tells us that when we kill someone who has one month left to live what we do is as wrong as killing someone with fifty good years to live. So along with the other two premises, it implies that killing someone who had fifty good years left to live is as wrong as knocking someone unconscious for a month in the middle of her life. Because that result is absurd, we should reject the Equal Wrongness Thesis.

In the next section, I clarify the Equal Wrongness Thesis. In section 2, I explain how my strategy connects that of a related argument developed by Kasper Lippert-Rasmussen,[5] and give some reasons for thinking his argument is ultimately unsuccessful. In sections 3 and 4, I defend the two premises of the main argument. I conclude in section 5.

1. Preliminary remarks

Several initial clarifications of the Equal Wrongness Thesis are necessary. First, as McMahan explains, the sense of "wrongness" that is

[5] Lippert-Rasmussen 2007

relevant for the thesis is one that admits of degrees.[6] One act is more wrong than another, in this sense, insofar as it is subject to a moral objection that would be harder to outweigh by countervailing moral considerations. It is important to note, however, that the Equal Wrongness Thesis is intended to apply only to individual acts of killing considered on their own. Suppose that on Monday I kill a young person and on Tuesday I kill an old person. The Equal Wrongness Thesis tells us that what I do on each day is equally wrong, and would have been equally difficult to justify. But two acts being equally hard to justify on their own does not imply that in a choice between them we ought to be indifferent which we choose. Suppose we must choose *between* killing a young person and killing an old person.[7] The Equal Wrongness Thesis is meant to be consistent with the judgment that in this case we ought to kill the old person.[8]

Second, following McMahan and others, I will understand the Equal Wrongness Thesis to be restricted in scope to a certain class of killings.[9] It does not apply to killings that differ in morally relevant but extrinsic ways, such as their effects (positive or negative) on third-parties, or the pre-existing relationships between the perpetrator and victim. It may be more wrong to kill a close friend than a stranger. Nor does the thesis apply to killings that differ with respect to the presence or absence of potentially justifying considerations like the victim's liability to suffer defensive harm or her free consent to being killed. It does not apply to killings that differ in their modes of agency—whether the victim is killed as a means to some end or as a side effect of pursuing an end, for instance. Lastly, the Equal Wrongness Thesis concerns only the killing of those individuals who satisfy the conditions of personhood, whatever those conditions are. It is therefore compatible with any view about the wrongness of killing non-persons. In the rest of this article, I will usually

---

[6] McMahan 2002, p. 190. See also Lippert-Rasmussen 2007, p. 717. Cf. Soto 2013, pp. 551–4.

[7] Because, we might suppose, one or the other must be killed as a necessary side effect of averting the deaths of thousands of other innocent people.

[8] McMahan 2002, p. 237.

[9] McMahan 2002, pp. 236–7. Other discussants of the Equal Wrongness Thesis also restrict it in this way—see Lippert-Rasmussen 2007, pp. 718–19; Soto 2013, p. 544; Hanser 2013, pp. 392–409.

omit the preceding qualifications, taking them to hold implicitly in the instances of killing I consider.

2. Lippert-Rasmussen's trilemma

In "Why Killing Some People Is More Seriously Wrong than Killing Others", Kasper Lippert-Rasmussen also challenges the Equal Wrongness Thesis on the basis of an analogy between killing people and knocking people unconscious. Our substantive defenses of that analogy are very different, not least because, as I will argue in due course, I believe his defense of it is unsuccessful. But my overarching strategy in this article was inspired by his. My main argument begins in the next section, in which I defend its first premise. Here, I want to explicate the structure of Lippert-Rasmussen's argument, give a few reasons for thinking his approach is ultimately wanting, and situate his strategy in relation to my own.

Lippert-Rasmussen's challenge is based on a trilemma consisting of the following three claims[10]:

> (1) *The Equal Wrongness Thesis*: The degree to which different killings of persons are wrong does not vary: all such killings are equally wrong.
>
> (2) *The Unequal Wrongness of Renderings Unconscious Thesis*: For any period ω, it is more wrong to render a person unconscious for a period longer than ω than it is to render a person unconscious for ω, other things being the same.
>
> (3) *The Equivalence Thesis*: It is neither more, nor less, wrong to deprive a person of a certain amount of conscious experience by killing her than it is to deprive her of the same amount of conscious experience by rendering her unconscious, other things being the same.

To see that these three claims cannot all be true, consider the following two pairs of acts. In the first pair, I kill someone with fifty years left to live and kill someone else with just one year left to live. Ac-

[10] Lippert-Rasmussen 2007, p. 716.

cording to (1), these two acts are equally wrong. In the second pair, everything is the same but instead of killing them I administer to each a powerful drug that induces a total coma that will last until her death. According to (2), now what I do to the person who would live another fifty years is more seriously wrong than what I do to the person who would live just another year. If the preceding judgments are correct, then (3) cannot be true. For if depriving someone of a certain amount of conscious experience by killing her is no more or less wrong than doing so by knocking her unconscious, the relative wrongness of the acts would be the same in each pair. Lippert-Rasmussen argues that of these three claims we should give up the Equal Wrongness Thesis.

I take the crucial premise of this argument to be (3). There are a number of ways to challenge that premise. Chief among them: unlike when she is knocked unconscious, when a person is killed she will never regain consciousness. Since it may matter for the wrongness of an act whether it deprives its victim of her last period of conscious life, we might doubt that depriving someone of a given amount of conscious experience by knocking her unconscious really is morally equivalent to doing so by killing her.

Lippert-Rasmussen suggests two individually sufficient ways of responding to this worry. The first is to restrict (3) so that it covers only instances of unconsciousness that persist until the victim's death.[11] (Whether we take this restriction to be covered by a suitable interpretation of the premise's *ceteris paribus* clause, or else to constitute a revised version of the premise itself, is not a substantive matter.) If we take this route, (3) can no longer be challenged on the ground that it makes a moral difference whether or not the victim will regain consciousness. The other response is to keep the more expansive, unrestricted version of (3), and argue independently that it is not in itself morally significant whether an act deprives its victim of what would have been her last period of consciousness. I begin by discussing the latter strategy.

Lippert-Rasmussen proposes that it only seems more wrongful to deprive someone of what would have been her last conscious period because we mistakenly assume either that it would be an especially great loss for her, or that it is especially disrespectful to deprive her

[11] Ibid., p. 727.

of it.[12] With respect to whether it would constitute an especially great loss, Lippert-Rasmussen notes that there is no reason, in principle, as to why the final period of a life is more valuable than any other period.[13] Yet even if we think the wrongfulness of an action tracks its harmfulness in this way, that overlooks the possibility that it is in itself bad for a person to suffer the permanent loss of conscious life.[14] As for the *disrespectfulness* of depriving someone of her last period of consciousness, it is not far-fetched to think that ensuring that someone will never again experience anything at all is a particularly egregious offence in its own right. Before we can accept the more expansive version of (3), we need to do more to show that the wrongness of knocking someone unconscious does not depend on when in her life the unconsciousness occurs.

In light of these issues, perhaps the opponent of the Equal Wrongness Thesis should adopt the restricted version of (3)—the version which only compares killing to cases of knocking someone unconscious in which the unconsciousness persists until the the victim's death. I myself argue that the restricted version of (3) is correct. But in the context of Lippert-Rasmussen's own argument, restricting (3) in this way is problematic.

For the trilemma to go through with this restriction, (2) must also be be understood to apply only when the instances of knocking people unconscious under comparison each last for the rest the victims' lives. Lippert-Rasmussen states early on in his argument that he will not entertain rejecting (2).[15] But it seems to me that we may have found (2) compelling in the first place because of the implicit assumption that the cases of knocking unconscious it compares do not both end the victim's conscious life forever. The trilemma is forceful precisely because of the independent plausibility of the Equal Wrongness Thesis, after all—and those who initially thought that the wrongness of killing someone does not vary according to how much life is taken from her may not find it obvious that the wrongness of causing someone to be *permanently* unconscious depends on its duration.

---

[12] Ibid., p. 726.

[13] Ibid.

[14] Kamm 1993, pp. 30–31; 2017; Kagan 2012, pp. 278–9; Benatar 2017, pp. 102–110.

[15] Lippert-Rasmussen 2007, p. 722.

To defeat the Equal Wrongness Thesis, we need to show both that the wrongness of knocking someone unconscious does not depend on when in the victim's life that unconsciousness occurs, *and* that any other features of killing that could plausibly distinguish it from knocking someone unconscious—that killing someone causes her nonexistence, for example—are not in fact morally relevant. The rest of this paper is an attempt to do just that. My strategy is to first defend, in a different way, the restricted version of (3). I then offer an argument that the wrongness of knocking someone unconscious for a given period of time is unaffected by when in her life it occurs. Together, these two premises should be enough to refute the Equal Wrongness Thesis, since along with that thesis, they imply that killing someone a month before she would otherwise have died is as wrong as killing someone with fifty good years left to live.

3. The moral equivalence of killing someone and knocking someone unconscious until her death

In this section, I will defend the argument's first premise:

> (A) It is as wrong to deprive someone of a certain amount of conscious life by knocking her unconscious until her death as it is to deprive someone of the same amount of conscious life by killing her, other things being the same.

In his own defence of that claim, Lippert-Rasmussen employs a pair of cases, in which

> (1) I kill a person twenty years before he would otherwise have naturally died;
>
> (2) I knock a relevantly similar person unconscious for twenty years, at the end of which he dies a natural death, and I do so knowing that he will not regain consciousness before dying.[16]

[16] Ibid., p. 722.

Lippert-Rasmussen proposes, first, that these two acts are equally bad for their victims.[17] He acknowledges that the fact that two acts are equally bad for their victims does not imply that they are equally wrong. (Indeed, if an act's wrongness always with its harmfulness we could reject the Equal Wrongness Thesis out of hand.) He argues, however, that in the absence of an explanation for why two acts differ morally, the fact that they are equally harmful is a reason to believe them to be equally wrong.[18] But if two acts *seem* to differ morally, then even lacking an explanation for why they differ, we are probably unwarranted in concluding them to be equally wrongful on the basis that they are equally harmful—and it is questionable that most people would pre-theoretically believe (1) and (2) to be equally wrong. Moreover, in fact there are a number of reasonable candidate explanations for that difference: for example, that killing someone ends that person's life, whereas knocking her unconscious does not.

The other argument Lippert-Rasmussen gives for the moral equivalence of killing someone and knocking someone unconscious until her death is that this equivalence follows from existing accounts of *why* killing is wrong.[19] If the factors that explain why killing someone is wrong are present in equal measure when knocking someone unconscious, it is reasonable to assume that the two types of act are morally equivalent. For instance, if the intrinsic wrongness of killing was wholly explained by its harmfulness, then on the plausible assumption that being knocked permanently unconscious is no better or worse for a person than being dead, we could conclude that the strength of the moral objection to the two types of act was the same. Alternatively, suppose the wrongness of killing is fully explained by its disrespectfulness to the victim. Then, on the assumption that the factors that make killing someone disre-

[17] Ibid. He adds that any cost it would be permissible to impose upon me to prevent my doing (1) could also be permissibly imposed upon him to prevent him doing (2). I agree with Matthew Hanser (2013, p. 379n17) that this argument would only sway someone to whom these acts already seem morally equivalent.

[18] Lippert-Rasmussen 2007, p. 722.

[19] Ibid., pp. 723–4.

spectful apply equally to knocking her unconscious, quite probably the two types of act are morally equivalent.[20]

The general problem with this argument is that whether or not all killings are equally wrong is itself a major consideration that any account of why killing is wrong must accommodate. Indeed, it is primarily on the basis of his conviction that all killings are equally wrong that McMahan rejects accounts of the wrongness of killing that imply otherwise, and it is also on the basis of that conviction that he proposes his own Intrinsic Worth Account, whose chief selling point is that it entails the Equal Wrongness Thesis.[21] If knowing whether the Equal Wrongness Thesis is true is necessary to adjudicate among competing theoretical accounts of the wrongness of killing, we are probably unwarranted in using those accounts to decide whether the Equal Wrongness Thesis is true.

In the rest of this section, I will attempt to establish the moral equivalence of killing and knocking unconscious until death in a different way. My argument runs as follows. If, other things being equal, killing someone really is more wrongful than knocking someone unconscious until her death, then killing someone must have some wrong-making property not shared by knocking someone unconscious until her death. But there is no such property. Therefore, the two acts are morally equivalent.

I take it that the conceptual difference between killing someone and knocking her unconscious until her death is that the former causes the victim's death, whereas the latter merely ends her conscious mental activity until she dies in a causally unrelated way. Thus, how we should conceptually distinguish the two kinds of act will depend on how we define death. Though numerous definitions of death (and by extension of killing) have been offered, each falls

---

[20] Ibid., p. 724. In defence of the claim that knocking someone unconscious is *ceteris paribus* no less disrespectful than killing her, Lippert-Rasmussen points to the fact that persons deserve respect not in virtue of their simply being alive, but rather in virtue of "certain cognitive and emotional capacities" (p. 724). However, an unconscious person might retain her capacities for cognition and emotion. Furthermore, the argument seems to conflate the *bases* of respect for persons and the *manner* in which we should respect them. That we should respect people in virtue of certain capacities they possess does not imply that respect for them consists in respect for those capacities.

[21] McMahan 2002, pp. 241–3.

into one of two categories. The first category, which we can call the *Existence Conception*, identifies a person's death with the cessation of her existence. If death is the cessation of existence, then the difference between killing someone and knocking her unconscious until her death is roughly that the former ends the person's existence, while the latter merely causes her to have no conscious mental activity until her existence ends for an unrelated reason. As we shall shortly see, as a conception of killing the Existence Conception needs fine-tuning, but this rough characterisation will do for the moment.

The second category of definitions of death, which we can call the *Biological Conception*, identifies a person's death with the cessation of functioning of one or more of her basic organismic processes. On the Biological Conception of killing, then, the difference between killing a person and knocking her unconscious until her death is that the former causes the cessation of the vital biological processes while the latter merely causes her to have no conscious mental activity up to the time that those processes cease due to independent causes.

These two conceptions—Existence and Biological—encompass the plausible definitions of death. Hence, they encompass the plausible views about what could differentiate killing a person from nonlethally depriving her of all future conscious experiences. Therefore, if it can be shown that on neither of these two conceptions of death is causing death a factor in determining the degree to which an act is wrong, it will have been shown that the two types of act are morally equivalent.

I begin with the Existence Conception. On this conception, the difference between killing someone and knocking someone unconscious until her death is that while the former brings about her nonexistence, the latter merely causes her conscious mental life to cease up to the time that her existence ends for other reasons. Matthew Hanser is adopting this conception of killing when he writes that killing a person is more seriously wrong than knocking her unconscious because "[s]omeone who has been . . . knocked unconscious continues on in an impaired state; someone who has been killed does not continue on at all."[22]

This point recalls one made by Frances Kamm. She writes:

---

[22] Hanser 2013, p. 399.

> Suppose we put someone into a coma, knowing that he will never recover from it . . . In such a case, we would not, I believe, be treated as harshly as if we had killed someone . . . The fact that one person determines the nonexistence of another against his will . . . is a factor in making killing wrong.[23]

For the purposes of moral analysis, construing killing as causing someone's nonexistence is probably strictly speaking a mistake. For there are well-known cases in the theory of personal identity which arguably show that an individual can be caused to cease to exist in a way that seems morally quite unlike killing her.

Consider first what looks to many people like a case of wrongful killing:

> *Riding Roughshod*: I am driving to the hospital to receive an urgent life-saving treatment when I notice you crossing the road ahead of me. Pressed for time, I run you over, killing you.[24]

Contrast that case with the following one:

> *Your Involuntary Division*: I am driving to the hospital to receive an urgent life-saving treatment when I notice you crossing the road ahead of me. I press a button on my dashboard that causes you to divide, amoeba-like, into two individuals, and then deposits each of your two "offspring" on either side of the road, where my car passes safely between them.

Like in *Riding Roughshod*, arguably pressing the button in *Your Involuntary Division* causes your nonexistence. Since you cannot be numerically identical with both resulting people, you will not be numerically identical with any individual who exists afterward.[25] Yet to press this button is plainly not as objectionable as fatally running you over. My own intuitions are that causing your nonexistence is

---

[23] Kamm 1993, p. 21.

[24] See Thomson 1991, p. 290.

[25] See Parfit 1984, pp. 255–260.

impermissible in *Riding Roughshod* but permissible in *Your Involuntary Division*. But whatever one thinks about the overall permissibility of either case, it is clear enough that the moral objection to my action is much weaker in the second case than in the first.

It might be replied that unlike causing your nonexistence by running you over with my car, causing your nonexistence by splitting you in two preserves, or mostly preserves, "what matters in survival".[26] Perhaps that is why it seems less objectionable to kill you in *Your Involuntary Division*. Since this reply presupposes that the moral objection to causing a person's nonexistence depends on its badness for the victim, to give it is already to reject the Equal Wrongness Thesis.

I do not deny that there is some moral objection to causing a person to divide in two against her will. But it seems to me that whatever objection there is, it does not arise from the bare fact the doing so brings about her nonexistence.[27]

The preceding remarks should not lead us to abandon the Existence Conception wholesale. Instead, we should accept that there are two salient ways of causing an individual's nonexistence, which we might call "killing by annihilation" and "killing by division". Killing by annihilation is a matter of causing something to cease to exist by eliminating its essential *intrinsic* properties (those intrinsic properties which it could not possibly lack). Killing by division is a matter of causing something to cease to exist by eliminating its essential *extrinsic* properties—for example, by creating one or more continuers of it, none of which is numerically identical to the original. The version of the Existence Conception that is useful for moral theorising about killing is killing by annihilation.

With this in hand, recall that (A) claims that depriving someone of a given amount of conscious life by knocking her unconscious until her death is as wrong as depriving her of the same amount of conscious life by killing her. Since the version of the Existence Conception of killing that we care about concerns eliminating the essential intrinsic properties of someone, to assess (A), we need to ask what intrinsic properties of people are essential to them. I shall ar-

---

[26] Parfit 1971; 1984, pp. 253–66; 1995, pp. 41–3.

[27] It arises, I submit, from the fact that it involves the violation of the victim's bodily integrity or some cognate consideration concerning her moral rights against interference.

gue that on no mainstream account of those properties is there a moral difference between causing someone's death by eliminating that property, on the one hand, and knocking her unconscious until she dies for unrelated reasons, on the other.

Consider first, then, the psychological continuity account of persons, according to which a person will exist in the future if and only if there will be someone to whom she is (uniquely) connected via overlapping chains of beliefs, memories, desires that lead to actions, and so on.[28] Now, if the relevant psychological states are taken to be conscious ones, then there is quite literally no difference between the two types of act: depriving someone of all future conscious experience just *is* killing her. Because most of our psychology can exist without being consciously active, however, on what is perhaps a more plausible version of the psychological continuity condition it is possible to knock someone unconscious until her death without causing her to cease to exist. On this condition for continued existence, the difference between killing someone and knocking her unconscious until her death is that the latter does not terminate the psychological states or processes that ground her persistence over time.

Alongside the psychological continuity view, consider what McMahan has called the "embodied mind account".[29] According to this account, a person continues to exist if and only if there is the unique physical and functional (or potentially functional) continuity of enough of those parts of the brain to retain the capacity for consciousness. On the embodied mind account, the difference between killing someone by causing her nonexistence and merely knocking her unconscious until her death is that only the killing involves the destruction of enough of the relevant parts of the brain.

I believe that on neither the psychological continuity account, nor on the embodied mind account, is it in itself morally relevant whether or not one's action causes someone's nonexistence. To see why, it will be instructive to consider two real ways a human being can be biologically alive even though she will never again be conscious.[30] The first way is a permanent vegetative state, in which the

---

[28] Shoemaker 1970; 1984; Perry 1972; Lewis 1976; Parfit 1984, 204–209.

[29] McMahan 2002, 66–94. Later in his career Parfit (2012) defended a similar view.

[30] For illuminating surveys of these phenomena, see McMahan 2002 (especially pp. 423–455) and Horne 2009.

victim has permanently lost consciousness because of irreversible damage to her cortex and limbic system. In a permanent vegetative state, the only parts of the brain that retain any functional capacity are the brainstem's vegetative centres: those responsible for heart rate, blood pressure control, temperature control, and respiration. The other way involves damage to the ascending reticular activating system, a network of cells in the brainstem that is necessary for any consciousness to occur.[31] A patient with localised damage to the reticular activating system retains the *capacity* for consciousness, as well as the bulk of her distinctive psychology, which is stored in (or constituted by) parts of the cerebrum. But because of that damage, as a matter of fact she will never again have any conscious experience.

Suppose that the conditions for a present individual to continue existing into the future are either the continuity of her psychological makeup or else the physical and functional continuity of those parts of the brain that have the capacity for consciousness. If it was in itself morally objectionable to terminate a person's existence (keeping in mind that we are using the "annihilation" sense of doing so), then it would be in itself morally objectionable to end a patient's life by terminating life support if she was in a total coma from which she would never recover due to damage to the reticular activating system, but not to remove life support from a patient in a permanent vegetative state. That is because the former act would bring about the discontinuity of the patient's psychology as well as end her capacity for consciousness, whereas for a patient in a permanent vegetative state it would not, for these phenomena have already ceased. It is very hard to believe that there is this moral difference.

The preceding argument indicates, I believe, that on neither of the two accounts of the conditions for continued existence that we are considering is it in itself morally objectionable to cause a person's nonexistence. But the argument shows this by appealing to atypical cases in which the conscious life of the victim has already in fact ended forever. It might be objected that terminating a person's existence *is* independently objectionable, but only so long as the conscious life of the victim is not already over. According this suggestion, the wrongness of ending a person's existence exhibits a "combination effect": when ending a person's existence *also* ends her

[31] Zeman 2001; McMahan 2002, 85–86, 441–43; Horne 2009.

conscious life, the former constitutes an additional moral objection over and above the wrongness of the latter taken in isolation.

It would not, however, be more wrongful to cause a conscious person to be in a permanent vegetative state than it would be to cause her to be permanently unconscious by destroying her reticular activating system. That is so even though only the former would involve the termination of psychological continuity or the destruction of those parts of the brain that have or provide the capacity for consciousness. It is not plausible that there is a stronger moral objection to ending someone's conscious life forever based only on which part of the brain one damages. The capacity for consciousness is not meaningful if nothing can exercise that capacity. And although the psychological makeup of a patient with permanent damage to her reticular activating system arguably persists, it is only "stored". The fact that a person's *non*-conscious psychology is not destroyed would not, by itself, mitigate the wrongness of terminating her conscious life.

It remains to consider the biological continuity condition of existence, according to which a person will exist in the future if and only if the organism that she is continues to exist and function biologically.[32] If continued existence consists in biological continuity, however, then being killed on the Existence Conception of death is coextensive with being killed on the Biological Conception. Because I will now argue that killing someone and knocking someone unconscious until death are morally equivalent when killing is construed according to the Biological Conception, I can establish that the two types of act are morally equivalent on the biological version of the Existence Conception incidentally, since the arguments will also apply to it.

Let us turn, then, to the Biological Conception of death, according to which the difference between killing someone and knocking her unconscious until her death is that the former act but not the latter involves the termination of the victim's vital biological processes. In other passages, Hanser seems to be adopting this conception of killing instead of the Existence Conception. He writes:

> An agent arguably interferes with someone in a much more fundamental way by killing him than he does by knocking

[32] Olson 1997; Snowdon 2014.

> him unconscious . . . The former causes the complete and permanent cessation of the victim's basic life-sustaining (or life-constituting) bodily operations; the latter leaves the victim alive but, for a time, unable to exercise a variety of his practical and perceptual capacities.[33]

Hanser here cites two distinct features of killing someone which allegedly make doing so more wrong than knocking someone unconscious. One is that it causes the cessation of the victim's vital bodily operations; the other is that this cessation is permanent. These components are separable, at least in principle. Conceivably, medical advances might one day allow for the biological resuscitation of a body that has ceased to function completely.

We should ask which of the two features of killing (or both) is morally significant. Consider first the appeal to the permanence of killing. Assuming that one's victim is mortal, killing her does not make it the case that she will be permanently dead; rather, it hastens her permanent death. So, if the permanence of death is a factor in making killing wrong, that must be because in killing a person one becomes the *cause* of her permanent death.

To see whether this could matter morally, contrast a pair of cases in which, absent any interference, the person I kill would have lived forever:

> (1) I cause my victim's vital biological operations to permanently cease at t_1, thereby pre-empting their permanent cessation at t_2.
>
> (2) I cause my victim's vital biological operations to cease for the interval t_1–t_2. At t_2, they will be made to cease *permanently* by some independent process (one I fully foresee but will not cause). This other process would have occurred when it did whether or not I had interfered with my victim first.

It seems to me that what I do in (1) is no more wrong than what I do in (2), notwithstanding the fact that in only (1) do I *cause* the victim's permanent biological death. This indicates that causing the temporary cessation of someone's biological functions which will thereafter cease permanently for other reasons is no less wrong than

[33] Hanser 2013, p. 399.

causing their permanent cessation. As further support of this claim, consider an analogous pair of cases in which the harm for which I am responsible is non-fatal:

> (3) Albert will be infected by the incurable debilitating Virus X next week if and only if I don't pre-emptively infect him with it today.
>
> (4) Bernard will be infected by Virus Y next week whether or not I infect him with Virus X today. Virus Y has identical symptoms to Virus X, but it will also completely inhibit Virus X if that virus is present in the host when Virus Y is contracted.

What I do to Albert in (3) is not, I believe, more seriously objectionable than what I do to Bernard in (4). Some support for this claim may come from the observation that if I had to infect just one of these men, it seems that I would have no more reason to choose Bernard. That is so even though only in Albert's case would I be responsible for the event that is someone becoming debilitatingly ill for the rest of his life. Suppose, furthermore, that we re-imagine the cases so that the viruses are beneficial, rather than harmful, to their hosts. It seems clear that I would have no more moral reason to infect Albert with the beneficial Virus X*, on the assumption that he would anyway contract it next week if I don't do so now, than I would have to give X* to Bernard, who will contract the superseding but equally beneficial Y* next week no matter what I do. Assuming that benefits and harms are analogous in the relevant way, this supports the claim that, in instances of non-lethal harming, the wrongness of one's act is determined by the difference one makes to what happens to the victim, not by the effect on her for which one is responsible. Without a good reason not to, it seems to me that we should generalise this result to lethal harms. The mere fact that killing someone causes the permanent cessation of her vital biological operations is not, I conclude, a factor in making it wrong.

Instead of appealing to the permanence of death, one might try to show that killing is especially wrong by appealing to its *irreversibility*. It might be said that depriving someone of conscious life by killing her is more objectionable than doing so by knocking her unconscious because the effect of the former on the victim cannot be

undone. We can understand this claim to be about metaphysical, physical (i.e., nomological), or practical irreversibility. Consider first the proposal that killing is especially wrongful because it is either metaphysically or physically irreversible. Suppose I wilfully terminate all of a person's mental processes along with her biological life. Her condition is not metaphysically or physically irreversible, let us suppose, for I know that the drug that would revitalise her has been successfully manufactured in the past. However, I also know that all samples of the drug have been destroyed and the knowledge of how to manufacture it has been lost forever. It is both metaphysically and physically possible to make the drug again, so my victim's condition is reversible in both those senses—it is just that I know for a fact that it will not be reversed. I believe that it would be no less wrong to end my victim's life in this case than it would be if the drug to revitalise her had never existed and indeed never could have existed.

Consider next the claim that killing is especially wrong when and because the death it causes is irreversible in a practical sense. Killing a person, it might be said, is objectionable in part because by doing so one makes it the case that there is no feasible way to bring that person back to life. Against this claim, first suppose the revitalising drug exists, but the last remaining sample is owned by an independently acting agent who I know will never use it to restore the life of my would-be victim. Assume that I know, too, that there is nothing I can do to change this. When I end my victim's life in this scenario, my effect is reversible in the practical sense; there is a real way to bring her back to life. Yet it seems to me that it would be no less wrongful to end her life in that scenario than it would be if the drug had never existed. Because this argument appeals to what I and another agent bring about, it may seem misleading to focus on the wrongness of my action alone. But the mere fact that there is someone *else* who could revive my victim does not mitigate the wrongness of my own action in killing her, given that I know, when I act, that this other person will never in fact revive her and cannot be made to do so. Ending her life would be no less wrongful if the reason the drug could not be used was some impassable natural obstacle instead of an agential one.

I have argued that neither the permanence of death nor its irreversibility is a factor in making killing a person more wrongful than knocking someone unconscious until her death. Let us turn, then, to the more salient potential wrong-making feature of the Biological

Conception: simply that it brings about the cessation of the victim's life-sustaining (or life-constituting) bodily operations. I believe that the fact that one act causes the cessation of a person's bodily operations does not, on its own, make its performance more objectionable than another. Suppose that someone in a temporary and medically non-serious coma will soon be connected to a life support system. While she is connected to the machine, if the parts of her brain that support her mental life are destroyed, her bodily operations will continue to function: her heart will pump blood, her lungs will inhale and exhale, and so forth. If those same parts of her brain are destroyed while she is *not* connected to the machine, however, the injury will also cause her biological death. If bringing about the cessation of a person's vital biological operations were in itself wrongful, then it would be less objectionable to obliterate this person's brain while she was on the life support system than to do so before she was connected to it, for in the former case her biological operations would continue to function even after the totality of her mental life was gone. But it is very difficult to believe that there is a fundamental moral difference between these two acts.

Earlier in this section, I postponed arguing that it is not in itself wrongful to terminate a person's existence if personal persistence is a matter of biological continuity. I take the preceding arguments concerning the Biological Conception of killing to have now established that conclusion, too. Hence, on none of the mainstream accounts of the conditions for our continuity is terminating someone's existence a factor in making killing wrong.

To summarise, on neither the Existence Conception nor the Biological Conception of killing is there is an intrinsic moral difference between killing a person and knocking her unconscious until her death. Because these are the only two tenable conceptions of killing, and because one type of act is more wrong than another only if there is some morally significant property of the former that is not shared by the latter, I conclude that it is just as wrong to knock someone unconscious until her death as it is to deprive her of the same amount of conscious life by killing her.

4. The moral irrelevance of when in her life a person is knocked unconscious

I turn now to the second premise in my argument against the Equal Wrongness Thesis:

> (B) It is as wrong to deprive someone of a certain amount of conscious life by knocking her temporarily unconscious in the middle of her life as it is to deprive someone of the same amount of conscious life by knocking her unconscious until her death, other things being the same.

Before proceeding, it is important to clarify two things about (B)'s *ceteris paribus* clause. First, it should be understood to exclude any differences in the prudential value for the victim of the conscious period of which she is deprived. It may be said that the last month of a person's life would have been especially important because it would have allowed her to say her goodbyes, or to realise some end toward which she had been working. But these are contingent facts about a life, and anyway I suspect that for many people who will live until old age a conscious month in the middle of their life is worth at least as much to them as the one at its very end. Second, (B) should be understood to exclude differences in the length of the victims' counterfactual (conscious) lifespans. It is plausible that it is more wrong to deprive someone of a given amount of consciousness the shorter her total conscious life would other have been, but the premise here applies only when that potential difference is factored out.

To get a better sense of (B), consider the following illustration. Calvin and Dylan will both live for exactly 1,000 months (about 83 years). Calvin has lived for 500 months so far while Dylan has lived for 999 months. Apart from their ages, the two men differ in no morally relevant respect. Suppose that the prudential value of what would be Dylan's 1,000th conscious month is the same as that of what would be Calvin's 501st month. (B) says that the strength of the moral objection to knocking out either man for one month is the same.

It might be claimed instead that the deprivation of his last conscious month is worse for Dylan than the deprivation of a conscious month in the middle of Calvin's life is for him. Because we are as-

suming that the prudential value of the month each man loses is the same, this is not plausible as a claim about lifetime wellbeing.Though it is true that Dylan has less remaining life, he has to that extent lived more life already.

Instead, it might be said that the deprivation is worse for Dylan in an absolute or non-comparative sense. Frances Kamm has argued that death is a non-comparatively bad event—bad in itself as opposed to bad in virtue of the goods it precludes—because death involves the loss of one's last period of conscious experience and therefore means that "everything is all over for one".31 The same is true when the deprivation of one's last period of conscious life is not due to death but to being placed in an unconscious state. If Kamm is right, then even if a conscious period at the end of a person's life would have contained no more value for him than one in the middle of his life, depriving him of the final period would be more harmful, because it also involves an additional, non-comparative harm. That might be enough to make knocking someone unconscious until his death more wrongful than knocking someone temporarily unconscious.

If the permanent loss of consciousness really is a non-comparatively bad thing to happen to someone, we should in principle be able to ask how bad it is for her. In the first place, however, it is utterly unclear how to go about assigning to a given death a non-arbitrary value such that it is equivalent, for example, to some degree and duration of pain. But let us grant that we can assign it a value, if only an imprecise one. It is still implausible that a the permanent end of a person's conscious life constitutes a non-comparative bad for her, for it is implausible that there is any significant amount of non-comparative *good* that it could outweigh. Suppose, for the sake of exposition, that death's non-comparative badness is such that to offset its presence in a person's life she would need to live for at least three happy years. Then a life of two happy years would not be worth living, a life of one year worse still, and so on. But I cannot believe that the life of a child who dies after two happy years was overall bad for her. If anything, most people would be glad for the child's own sake that she got to live at all.

Of course, perhaps death is just not as bad as all that. Let us suppose it is only bad enough to outweigh three happy months. Would anyone conclude, upon the discovery of an extra-terrestrial species whose members die painlessly after only one merry summer, that

the total prudential goodness in the universe is lower than previously thought? Defenders of the position that death is non-comparatively bad for its subject might reply that the position applies to persons only, but things seem no different if we clarify that these extraterrestrials come into existence as fully-formed persons. For my part, it is hard to believe that death could be non-comparatively bad enough to outweigh even one happy day.

Suppose that this argument does not succeed; death really is non-comparatively bad. For all that, it is still not obvious that this could provide a prudential reason to postpone death. Were immortality possible, then choosing to live forever might be in one way good for me, since only then could I escape the non-comparative badness of death. But I must die eventually, and when I do my death will be no less bad for having been rescheduled. For death's non-comparative badness to justify postponing death in the Grim Reaper scenario, or for that matter in the Limbo Man case, it seems that it must also be true that it is in a person's interest to undergo a bad event later rather than sooner. That is a questionable premise which to my knowledge has not been adequately defended. We cannot, I conclude, appeal to the non-comparative badness of losing one's last period of consciousness in order to refute (B).

The other way to challenge the premise is to skirt questions of badness and proceed directly to the wrongness of the acts. In a response to Lippert-Rasmussen's article, Daniel Cohen and Morgan Luck have argued that the wrongness of depriving a victim of a given amount of consciousness is greater the nearer the victim is to the end of her conscious life.[34] More precisely, the authors defend the following principle:

> The wrongness of rendering someone unconscious corresponds with the proportion of the victim's *remaining* conscious life that is thereby removed, other things being the same.[35]

Their argument for that principle involves imagining two alien species. Members of the first species (call them the "short-livers") always live for just two years, while members of the second species (call them the "long-livers") live for 1,000 years. The authors imag-

[34] Cohen and Luck 2009, pp. 396–401.

[35] Ibid., p. 399 (emphasis added).

ine coming across a newborn member of each species.[36] As they point out, it seems more seriously wrongful to render the newborn short-liver unconscious for one year—half its remaining life—than to knock the newborn long-liver unconscious for a year, leaving it to experience another 999 years when it awakens.

Cohen and Luck's principle gets the right answer in this case, since the short-liver is deprived of a greater proportion of its remaining conscious life (one year of two) than the long-liver is of his (one year of a thousand). But there is a simple variation on their case that shows that their principle has an implausible implication. The principle they defend would imply that it would be no more wrong to knock unconscious a short-liver who has been alive for one year than a long-liver who has already lived for 999 years, for in either case one removes the entirety of the victim's remaining conscious life. But that is hard to believe. When we knock the short-liver unconscious we deprive it of half its total conscious life, whereas when we do so to a long-liver we reduce its conscious life by just one one-thousandth.

It might be thought that the *ceteris paribus* clause of Cohen and Luck's principle is meant to factor out differences in how much conscious life the victim has so far enjoyed. But the main conclusion of their article is that a person's age has no bearing on the wrongness of killing her, so their argument cannot go through if cases in which the victims differ in the amount of conscious life they have lived are not covered by their principle.

Cohen and Luck's newborn aliens case relies on the intuition that it matters, when knocking a person unconscious, how much future conscious life she loses out on. But as the variation I have just offered shows, it also seems to matter for its wrongness how much conscious life a person has so far enjoyed. To account for both factors requires a middle ground. One plausible candidate is that the wrongness of knocking a person unconscious for some period corresponds not with that period's proportion to the victim's remaining conscious life, but rather with its proportion to the total lifetime conscious experience she would otherwise have had.

That new principle would be:

[36] Assume that members of both species are born full persons.

> The wrongness of knocking someone unconscious corresponds with the proportion she loses of the *total* amount of conscious life she would otherwise have enjoyed, other things being the same.

This principle gets the right answer in both of the short- and long-lived alien cases. It also entails (B), which compares equally long deprivations of consciousness at different times within equally long conscious lives.

I now turn to a positive argument for (B). In it, Calvin and Dylan again make an appearance. Suppose that I possess two drugs: a *fast-acting drug* that immediately knocks its taker unconscious for one month, and a *delayed-acting drug* that has no effect until one month before its taker's death, when it will knock him unconscious for that final month.[37] Consider, first, the following case:

> *Case One*: I secretly slip the fast-acting drug into the coffee of the 999-month-old Dylan and the delayed-acting drug into the coffee of the 500-month-old Calvin.

It is not plausible that what I do to either man in Case One is more wrong than what I do to the other. I deprive each man of the very same amount of experience, at the very same location in his life, and in the very same manner. It is true that the effect of my action takes place immediately for Dylan but is delayed by nearly forty years for Calvin. But it seems to me that mere temporal distance between an act and its effect could not be morally significant in itself.

Next, consider:

> *Case Two*: Calvin has an identical twin, Caleb. I slip the delayed-acting drug into Calvin's coffee, and the fast-acting drug into Caleb's.

I do not think it is plausible that the acts compared in Case Two are wrong to different degrees, either. Note that to dispute that claim one cannot appeal to the wrongness of making it the case that

---

[37] McMahan considers an example involving fast- and slow-acting poisons—in his case causing either an immediate or a delayed death—which is the inspiration for these cases 2002, p. 248.

'everything is all over' for a victim, since here both victims do have conscious life in their future. Moreover, both men are deprived of the same proportion of their remaining conscious lives. The salient difference seems to be that only Calvin loses the final month of his conscious life. But since it is not the last month he has left, and since there is not a relevant sense in which being deprived of that month of consciousness constitutes a greater loss to him, it is hard to see why the mere fact that the month is at the end of his life should make its deprivation more wrongful.

Taken together, these two equivalences—that in Case One and that in Case Two—show that it is morally irrelevant whether, when I deprive someone of a month of conscious experience, I do so at the end of his life or its middle. If giving the fast-acting drug to the 999-month-old Dylan is morally equivalent to giving the delayed-acting drug to the 500-month-old Calvin (Case One), and that latter act is morally equivalent to giving the fast-acting drug to the 500-month-old Caleb (Case Two), then giving the fast-acting drug to the 999-month-old Dylan is morally equivalent to giving the 500-month-old Caleb the fast-acting drug. That result, of course, is the same as the case with which we began this section: knocking out Calvin now, or knocking out Dylan now.[38]

5. Conclusion

It is now possible to bring together the pieces of my argument. The first two premises were

> (A) It is as wrong to deprive someone of a certain amount of conscious life by knocking her unconscious until her death as it is to deprive someone of the same amount of conscious life by killing her, other things being the same,

and

[38] The same argument can be given for longer deprivations of consciousness than one month. But even this specific equivalence is sufficient to undermine the Equal Wrongness Thesis.

(B) It is as wrong to deprive someone of a certain amount of conscious life by knocking her temporarily unconscious in the middle of her life as it is to deprive someone of the same amount of conscious life by knocking her unconscious until her death, other things being the same.

These two premises entail

(C) It is as wrong to deprive someone of a certain amount of conscious life by knocking her temporarily unconscious in the middle of her life as it is to deprive her of the same amount of conscious life by killing her, other things being the same.

Establishing (C) is enough to defeat the Equal Wrongness Thesis. Suppose we take the period of unconsciousness in question to be one month. We would then hold that knocking someone temporarily unconscious for one month in the middle of her life is as wrong as killing someone who has one month left to live. The Equal Wrongness Thesis tells us that killing a person who has one month left to live is as wrong as killing a person with fifty good years to live. Hence, the conjunction of the Equal Wrongness Thesis and (C) entails that, other things being equal, knocking a person unconscious for a month in the middle of her life is just as wrong as killing a person with fifty good years ahead of her. This result is absurd. Whatever considerations might justify knocking someone out for a month in the middle of her life, those considerations cannot possibly be as stringent as the considerations required to justify killing a person with fifty good years left to live. On pain of absurdity, therefore, we should reject the Equal Wrongness Thesis.[39]

[39] For valuable feedback, I am grateful to Susanne Burri, Tomi Francis, David Kinney, Kasper Lippert-Rasmussen, Chris Marshall, Michal Masny, Jeff McMahan, Max Muir, Michael Otsuka, Bryan Roberts, Tom Rowe, Bastian Steuwer, and Alex Voorhoeve.

References

Benatar, David (2017) *The Human Predicament: A Candid Guide to Life's Biggest Questions*. Oxford University Press.

Bundesverfassungsgericht, *Judgment of the First Senate of 15 February 2006*. BvR 357/05, paras. 1–156. Available online: http://www.bverfg.de/e/rs20060215_1bvr035705en.html

Campbell, Tim and Jeff McMahan (2010) "Animalism and the Varieties of Conjoined Twinning." *Theoretical Medicine and Bioethics* 31(4), pp. 285–301.

Cohen, Daniel and Morgan Luck (2009) "Why a Victim's Age is Irrelevant When Assessing the Wrongness of Killing." *Journal of Applied Philosophy* 26, pp. 396–401.

Hanser, Matthew (2013) "The Wrongness of Killing and the Badness of Death." In Ben Bradley, Fred Feldman, and Jens Johansson (eds.) *The Oxford Handbook of Philosophy of Death*. Oxford University Press.

Hare, Caspar (2008) A puzzle about other-directed time-bias. *Australasian Journal of Philosophy*, 86(2), pp. 269–277.

Horne, Malcolm (2009) Are People in a Persistent Vegetative State Conscious? *Monash Bioethics Review*, 28(2), pp. 1–12.

Kagan, Shelly (2012) *Death*. Yale University Press.

Kamm, F. M. (1993) *Morality, Mortality: Volume I: Death and Whom to Save From It*. Oxford University Press.

— (1996) *Morality, Mortality: Volume II: Rights, Duties, and Status*. Oxford University Press.

— (2007) "Review Article: The Ethics of Killing: Problems at the Margins of Life." *The Philosophical Review*, 116(2), pp. 273–280.

— (2017) "The Purpose of My Death: Death, Dying, and Meaning." *Ethics* 127(3), pp. 733–761.

Lewis, David (1973) *Counterfactuals*. Princeton University Press.

— (1976) "Survival and Identity." In A. Forty (ed.) *The Identities of Persons,* pp. 17–40.

Lippert-Rasmussen, Kasper (2007) "Why Killing Some People is More Seriously Wrong than Killing Others." *Ethics* 117(4), pp. 716–38.

— (2005) *Deontology, Responsibility, and Equality*. Museum Tusculanum Press.

— (2001) "Two Puzzles for Deontologists: Life-Prolonging Killings and the Moral Symmetry Between Killing and Causing a Person to Be Unconscious." *The Journal of Ethics* 5(4), pp. 385–410.

McMahan, Jeff (1995) "Killing and Equality." *Utilitas* 7(1), pp. 1–29.

— (2002) *The Ethics of Killing: Problems at the Margins of Life*. Oxford University Press.

Nagel, Thomas (1970) "Death." *Noûs* 4(1), pp. 73–80.

Olson, Eric (1997) *The Human Animal: Personal Identity Without Psychology*. Oxford University Press.

Parfit, Derek (1984) *Reasons and Persons*. Clarendon Press.

— (2008) "Persons, Bodies, and Human Beings." In John Hawthorne, Dean Zimmerman, and Theodore Sider (eds.) *Contemporary Debates in Metaphysics*. Blackwell.

— (2012) "We Are Not Human Beings." *Philosophy* 87(1), pp. 5–28.

Perry, John (1972) "Can the Self Divide?" *Journal of Philosophy* 69, pp. 463–488

Shoemaker, Sydney (1970) "Persons and Their Pasts." *American Philosophical Quarterly* 7(4), pp. 269–85.

— (1984) "Personal Identity: A Materialist's Account." In S. Shoemaker and R. Swinburne (eds.) *Personal Identity: Great Debates in Philosophy,* pp. 67–132.

Snowdon, Paul (1991) "Personal Identity and Brain Transplants." In David Cockburn (ed.) *Human Beings*. Cambridge University Press.

Soto, Carlos (2013) "Killing, Wrongness, and Equality." *Philosophical Studies* 164(2), pp. 543–559.

Thomson, Judith Jarvis (1991) "Self-Defense." *Philosophy and Public Affairs* 20(4): 283–310.

van Inwagen, Peter (1990) *Material Beings*. Cornell University Press.

Youngner, Stuart, Arnold, Robert, and Shapiro, Renie (2002) *The Definition of Death: Contemporary Controversies*. Johns Hopkins University Press.

Zeman, Adam (2001) "Consciousness." *Brain* 124, pp. 1263-1289.