

Identity-Relative Paternalism is Internally Incoherent

Eli G. Schantz

Identity-Relative Paternalism, as defended by Dominic Wilkinson, holds that paternalistic intervention is justified to prevent an individual from doing to their future selves (where there are weakened prudential unity relations between the current and future self) what it would be justified to prevent them from doing to others.[1] Wilkinson, drawing on the work of Derek Parfit and others, defends the notion of Identity-Relative Paternalism from a series of objections. I argue here, however, that Wilkinson overlooks a significant problem for Identity-Relative Paternalism—namely, that it yields unactionable and self-contradictory results when applied to choices where both options present potential harms to future selves.

To illustrate this point, let us consider the following:

A is an acclaimed oboist, for whom the practice and performance of music has brought deep joy throughout their adult life. *A*, however, develops a condition which will not only cause the rapid loss of their manual dexterity and capacity to hear, but will also certainly be fatal within the year. Suppose further that Drug Z, if taken now, will slow the loss of *A*'s dexterity and hearing, but will hasten their demise by a number of months.

A considers the choice of whether or not to take the drug. *A* believes that they have had a long and joyful life, but would, all else equal, like to spend as much time as possible with their family. However, *A* also understands that, without the capacity to enjoy the music which has been central to their life, that those remaining months would likely be joyless in a meaningful way.

Taking this background, let us consider the first of *A*'s options: to prioritize retaining their ability to enjoy music, and, therefore, to take the drug. Within this branch of the choice, there are two potential future selves: call these **A1** and **A2**.

A1. *A* chooses to take the drug and continues to play and enjoy music over their final days. *A1* dies content with their choice.

A2. *A* chooses to take the drug. But as their death approaches, *A2* becomes increasingly regretful of their choice: they find that it was not the music which brought them joy, but the opportunity to share it with the people they care about. *A2*, prior to their death, believes that they should have prioritized spending more time with their loved ones.

Does Identity-Relative Paternalism justify intervention in this case? To provide an answer, we must first assess two further questions: (I) if *A* stands in a weakened prudential identity relation to either of their future selves, and (II) if it would be justified to prevent *A* from doing to others what they do to their future selves. If we answer both questions affirmatively, then Identity-Relative Paternalism justifies intervention.

In regard to (I): *A* prioritizes retaining their ability to enjoy music over spending more time with their family, but *A2* has the opposite priority, and, therefore, stands in a weakened prudential identity relation to *A*. In regard to (II): *A* brings about the early death of *A2* and prevents them from spending time with their family, and I take it that we would prevent *A* from doing this to another individual. Identity-Relative Paternalism, therefore, justifies intervention to prevent *A* from taking the drug.

But let us consider the other branch of the choice: to prioritize spending more time with their family, and, therefore, to *not* take the drug. Within this branch, there are two potential future selves: call these **A3** and **A4**.

A3. *A* chooses not to take the drug. Though they lose their ability to enjoy music, they find great joy in spending time with their friends and family in their final months. *A3* dies content with their choice.

A4. *A* chooses not to take the drug. As they lose their ability to enjoy music, however, they find their final days devoid of the joy which music brought them. *A4*, prior to their death, believes that they should have prioritized retaining their ability to enjoy music.

To determine if Identity-Relative Paternalism justifies intervention in this case, we return to (I) and (II). In regard to (I): *A* prioritizes spending more time with their family over retaining their ability to enjoy music, but *A4* has the opposite priority, and, therefore, stands in a weakened prudential identity relation to *A*. In regard to (II): *A* robs *A4* of the joy of music in the final months of their life, and I take it that we would prevent *A* from doing this to another individual. Identity-Relative Paternalism, therefore, justifies intervention to prevent *A* from *not* taking the drug.

But having arrived at this conclusion, we also arrive at a contradiction. We saw that Identity-Relative Paternalism justifies intervention to prevent *A* from taking the drug and causing harm to *A2*; but it also justifies intervention to prevent *A* from *not* taking the drug and causing harm to *A4*. In the case of *A*, then, Identity-Relative Paternalism yields an *unactionable* result: even someone who unquestioningly accepts Identity-Relative Paternalism is unable to resolve what should or should not be done in this case. But further, Identity-Relative Paternalism yields a *self-contradictory* result—and is, therefore, internally incoherent.

Such incoherence, however, is not limited to the case of *A*, as this argument is generalizable: in any case where options *X* and *Y* present potential harms to future selves, Identity-Relative Paternalism justifies intervention to (i) prevent the choice of *X* and force the choice of *Y*, while simultaneously justifying intervention to (ii) prevent the choice of *Y* and force the choice of *X*. And insofar as essentially all significant choices involve options which pose potential harms to future selves, Identity-Relative Paternalism yields unactionable and self-contradictory results in essentially all significant choices. For Identity-Relative Paternalism to be useful whatsoever, then, it must answer the challenge posed by the case of *A*.

Is such an answer accessible to the Identity-Relative Paternalist? I foresee two potential arguments which might be offered.

First, the Identity-Relative Paternalist may claim that the justification for paternalistic intervention can be indexed to the likelihood of the harm itself, the quantity of harm done, and the likelihood of value conversion: in the case of *A*, for example, one might claim that the harm to *A2* or *A4* is more likely or severe than the other. In such a case, our justification to prevent one choice might be stronger than our justification to prevent the other.

Such an argument, however, puts the Identity-Relative Paternalist in a nearly inescapable epistemological quagmire. It is nearly impossible, if possible at all, to predict how an individual's values will change in response to a major life event. Someone who overcomes a serious illness, for example, may come to value their independence more, or they may come to value the ways in which they are dependent on their loved ones; they might value their health more, or perhaps they will value it less, given the ways in which they saw their life had value even without it. The same goes for potential harms: it is not only incredibly difficult to predict *what* will occur, but also *how* what occurs will intersect with the complex values of an

individual. So, even if the justification for intervening is indexed to the likelihood or quantity of harm or the likelihood of value conversion, Identity-Relative Paternalism is *still* unable to render a decision on a case without knowledge of the future which borders on omniscience. Even though Wilkinson foresees epistemological challenges for Identity-Relative Paternalism, he vastly overestimates its ability to provide cogent judgements in everyday situations: without a reliable way to understand potential harms or value conversions, Identity-Relative Paternalism can't escape self-contradiction.

Second, the Identity-Relative Paternalist may claim that the case of *A* improperly conjoins two distinct cases, one where *A* prioritizes family, and one where *A* prioritizes music: since *A* can only prioritize one, Identity-Relative Paternalism shouldn't be expected to account for both. I would argue that this claim, however, misrepresents everyday moral decision-making. Put yourself in the situation of *A*: you are faced with a choice which requires you to give up one of two things which you deeply value. You will consider the pros and cons of each option, and in the end, you will make a choice which reflects your values as you understand them—and you, like *A*, may get it wrong. That's why the case of *A* is challenging: it *precisely reflects* the uncertainty inherent in our decision-making, and rather than representing two distinct cases, represents the ways in which we struggle at times to understand what we do or will value.

Naturally, the Identity-Relative Paternalist will present counterarguments to the points which are raised here. Nevertheless, the case of *A* presents a significant challenge to Identity-Relative Paternalism which must be solved. And until this challenge is satisfactorily met, Identity-Relative Paternalism will remain a theory of questionable practical and theoretical merit.

Bibliography

- 1 Wilkinson D. The harm principle, personal identity and identity-relative paternalism. *J Med Ethics* Epub ahead of print: 15 February 2023. doi: 10.1136/jme-2022-108418.