Philosophy and
Phenomenological Research

Philosophy and Phenomenological Research
Vol. XCVI No. 3, May 2018
doi: 10.1111/phpr.12325

Essentialism and the Nonidentity Problem

SHAMIK DASGUPTA
Princeton University

> "You have gone from non-being to being by one of these agreements which are the only ones to which I care to listen. You were thought of as possible, as certain, in the very moment when, in a love deeply sure of itself, a man and a woman wanted you to be."
> –André Breton, *Mad Love*

1. The Nonidentity Problem

Sometimes, one faces a decision that will affect not only the future distribution of welfare, but also the identities of the people over which the welfare is distributed. This can lead to a so-called "nonidentity problem". For example, consider the following case:

> A man and a woman would like to conceive a child as soon as possible, but are told by their doctor that the man has an illness that affects his sperm. The illness is curable, but it will take one month of therapy. If they conceive while he is ill, their child will be born with a condition that leads to a poor quality of life; a life worth living but only barely so. If they wait a month till he is better, their child will not have the condition and will live a much richer life. They have a choice between two acts: conceive immediately, or conceive in one month when the man is cured.[1]

Many have the sense that it would be morally wrong to conceive immediately, and that this is in part because it would be bad for the resulting child. But how can we account for this? Suppose they go ahead and conceive immediately, and give birth to a baby girl with the condition. Call the child "Xia". The puzzle is that if they had waited a month, they would have had a *different child*. So how could their decision to conceive immediately be bad for Xia? To the contrary: their conceiving immediately was necessary for Xia to exist and live her worthwhile life!

A little more precisely, the puzzle is that the following four claims are individually plausible but jointly inconsistent:

(1) Conceiving immediately was morally wrong.
(2) Conceiving immediately was morally wrong only if it was bad for Xia.

[1] I leave open exactly what the condition is and the particular way it affects quality of life. The details here will depend on one's specific views about what determines the overall welfare of a life.

(3) Conceiving immediately was bad for Xia only if Xia would have existed had they waited.

(4) Xia would not have existed had they waited.

Which claim should we reject? This is known as a nonidentity problem.

In this paper I develop a view on which we need reject none of these claims. On this view, (1)–(4) are not inconsistent after all: there is an equivocation running through the puzzle, with the name "Xia" denoting something different in (4) than it does in (2) and (3). More fully, the idea is that there are in fact many entities in the vicinity of the couple's child. We do not recognize them all in everyday thought, so they are easy to miss. But they are there nonetheless. Some of these entities make (4) true, but those entities have no moral significance: they do not matter when thinking about what to do. The entities that matter are different, and are the entities that (2) and (3) are true of, if true at all. English terms like "person", "child", and "baby girl" do not normally distinguish these entities. Nor does the name "Xia", given how it was introduced; hence the equivocation.

This view has been largely overlooked in the literature on the nonidentity problem.[2] This might be because the view becomes apparent only when one critically examines claim (4), yet the literature tends to take (4) for granted and then infers that one of the other claims must be rejected instead.

For example, some reject (1) and say that conceiving immediately was not morally wrong after all. This is not the plausible view that, having had Xia, the couple would love her dearly and have no regrets over their decision to conceive immediately.[3] Rather, it is the surprising view that it was not morally wrong of them, at the time, to conceive immediately. Still, proponents of the view argue that we are forced into this view, given the truth of (2), (3), and (4).[4]

Others reject (2). They recognize that it follows from what is sometimes called the "Person Affecting Principle", according to which an act is morally wrong only if it is bad *for some person*; in the case above, Xia is the obvious person. But they think that the lesson of the above case is that we should reject (2), and so reject the Person Affecting Principle. On this view, conceiving immediately was wrong not because it was bad for Xia, but for some other reason—perhaps because it did not maximize total welfare, for example.[5]

Yet others reject (3). They recognize that it follows from the "Counterfactual Comparative" view of bad-for, on which an act is bad for someone iff she is worse off than she would have been had the act not been performed. For, supposing that Xia would not exist had her parents waited, it follows that she is not worse off than she would have been had her parents waited; so it follows on the Counterfactual

---

[2] Wolf (2009) comes close to endorsing it. But his discussion conflates it with the so-called *de dicto* approach to the nonidentity problem, which is different. I discuss the *de dicto* approach in section 12.

[3] See Harman (2009) and Wallace (2013) for a discussion of the nature of regret (or lack thereof) in these kinds of cases, and its relation to moral evaluation.

[4] Heyd (1988) takes this approach. Roberts (1998, 2007, 2009) takes this approach for some nonidentity cases but not others; see footnote 10 for more details.

[5] Parfit (1984, chapter 16) argued that this is the correct response to the nonidentity problem, though in Chapter 18 he endorses a related principle called the "wide Person Affecting Principle". On one reading of Hare (2007), he takes a similar line. Freeman (1997) and Harris (1998) also reject (2).

Comparative view that conceiving immediately was not bad for her after all, just as (3) states. Thus, some think that the lesson of the nonidentity problem is that we should reject (3), and hence reject the Counterfactual Comparative view. Such theorists typically propose non-comparative accounts of why conceiving immediately was bad for Xia. For example, Shiffrin (1999) argues that it was bad for Xia insofar as it produced in her a condition that impedes the exercise of her agency; Harman (2004) argues that it was bad for her if it caused her "pain, early death, bodily damage or deformity" (p. 93); and Velleman (2008) argues for a view on which conceiving immediately was bad for Xia because it violated her "right to be born into good enough circumstances" (p. 275). The upshot on all these views is that conceiving immediately can be bad for Xia regardless of whether she is better or worse off than she would have been had her parents waited.[6]

These are the standard responses to the nonidentity problem. They are importantly different, but they all take (4) for granted and infer on that basis that one of (1), (2), or (3) must be false. If the view I develop in this paper is *coherent*—never mind whether it is *true*—this inference is invalid, since (1)–(4) can all be true together. My primary aim is to show that my view is indeed coherent, and hence show that accepting (4) does not commit one to rejecting (1), (2), or (3). To be clear, there may well be other reasons to reject (1), (2) or (3); my point is that it is a mistake to think that we must reject one of them *on the basis of (4)*. More generally, the point is that while nonidentity cases like the above have been widely used as a reason to reject the Person Affecting Principle or the Counterfactual Comparative view of bad-for, this line of reasoning is invalid: one can accept both principles even in the light of nonidentity cases.

My secondary aim is to present some reasons to think that the view I develop is true. While the standard approach takes the *metaphysical* claim in (4) for granted and infers that some *ethical* claim in (1)–(3) must be false, my approach goes in reverse: it uses our *ethical* beliefs about harm and wrongdoing as a guide to *metaphysical* conclusions about the natures of the entities that matter. In this respect, my approach here more resembles the approach widely adopted in the literature on personal identity over time, in which ethical judgments about responsibility and prudence are used as guides to the metaphysical persistence conditions of persons, not the other way round. I do not think that the arguments I present here are decisive, but I do think they show that the view I develop is worthy of serious consideration.

The logic of the nonidentity puzzle is not limited to the case described above. Consider the question of whether we should conserve the natural environment so that our descendants 300 years hence can live fruitful lives, or else deplete it and leave our descendants with lives that are barely worth living. Many believe that we should conserve, but the puzzle is how to account for this. For suppose we are selfish and do not conserve, and in 300 years time our descendants are indeed struggling. The puzzle is that if we had conserved, we would have had *different descendants*—conservation involves

[6] Woodward (1986) defends a related rights-based view. Others who reject (3) include McMahan (1981), Hanser (1990), Bykvist (2006), and Roberts (2007, 2009). Liberto (2014) also proposes a view in this vicinity, arguing that conceiving immediately is bad for Xia because it exploits her. Note that I am using "bad for" as a catch-all term to encompass harming the child, wronging the child, and so on. So, when Shiffrin says that an act *harms* a child, and Woodward says that it *wronged* the child, it follows that the act was "bad for" the child as I am using the term.

radical changes in public policy, with people pursuing careers they wouldn't otherwise had, thereby meeting mates they wouldn't otherwise meet, and so on. So, far from being bad for our descendants, our selfishness is a necessary condition for their worthwhile existence! Why, then, is it wrong to deplete the natural environment? The same kind of puzzle has been discussed in connection with a variety of other questions concerning disability, reparations, and other topics besides.[7]

Is the view I develop a coherent position in these other cases too? If so, is it true? I will say little about these other cases, focusing instead on the case of the couple described above. There are cases and cases, and I do not claim that all should be treated alike. Still, it should become plausible that the view I develop is a coherent position to take in all these cases. It should also become clear how my arguments would carry over to these other cases, but I will not evaluate whether they are equally compelling.

Thus, my aim is not to establish that the view I develop is the correct approach in all nonidentity cases. Nor, consequently, is my aim to find a general theory of population ethics. My aim is merely to argue that, at least in the case of the couple, the view is coherent (sections 2–8), and might even be true (sections 9–13).

2. Essentialism

I said that the view I want to develop becomes apparent only when one examines (4) in some detail, so let us turn to this. Suppose the couple conceive immediately and have Xia. Claim (4) says that if they had waited, they would have had a different child. This is typically assumed without argument. It certainly sounds true, to my ear at least. But perhaps it is also perceived to be a scientific truth based on the biology of human reproduction, and hence not open to philosophical critique. If so, that is a mistake: it rests on substantive metaphysical presuppositions. Any appeal to (4) in ethical debate therefore rests on these metaphysical presuppositions too—as the saying goes, there is no such thing as metaphysics-free ethics; there is only ethics whose metaphysical baggage has been taken on board without examination.[8] So let us examine the baggage.

Various things could have been different about me: I am now sitting, but I could have been standing. But some things about me could *not* have been different: I am a human being, and (plausibly) I could not have been a sea cucumber instead. Suppose that one thing that could not have been different about me is my origins—suppose that I could not have originated from a different sperm and egg. It follows that if my parents had waited a month before conceiving, they would not have had *me*, since their child would have had different origins. The idea behind (4) is that Xia has some property—perhaps her origins, perhaps something else—that could not have been different about her, and that if her parents had waited the resulting child would have lacked that property.

---

[7] For discussion the nonidentity puzzle as it arises in relation to the question of reparations, see Sher (1981, 2005), Thompson (2001), and Shiffrin (2009). For discussion of the puzzle at it arises in relation to disability theory, see Savulescu (2002) and Wasserman (2005).

[8] With apologies to Dan Dennett. The real quote, from "Darwin's Dangerous Idea", is: "There is no such things as philosophy-free science; there is only science whose philosophical baggage has been taken on board without examination".

This idea presupposes a view known as "essentialism". This is the view that there are two ways to have a property: *essentially*, and *accidentally*. As a working definition, let us say that x has P *essentially* iff having P is a necessary condition for being x. That is:

x has P *essentially* iff necessarily, for all y, $y = x$ only if y has P.

And say that x has P *accidentally* iff x has P, but not essentially.[9] I have the property of sitting, but only accidentally: it is possible for me to be standing. I also have the property of being human, but this time I have it essentially: nothing can be *me* without being human. The idea behind claim (4) is that Xia has a certain property P essentially, and that if her parents had waited a month the resulting child would not have P. It follows that the resulting child would not be Xia, as (4) states.

What might property P be? One could appeal to Xia's origins, her genetics, her date of conception, as well as other more complex properties. For now, the choice does not matter. This is not to say that the choice is entirely inert. If one says that her date of conception is essential, then (4) is true because it is *impossible* for the couple to wait and conceive Xia. By contrast, if one says that her genetic code is essential, then strictly speaking it remains *possible* for the couple to conceive a month later and have a child with that genetic code who is Xia; the idea would be that (4) remains true because that is so *unlikely*. I will return to this distinction later on, but for now it does not matter: whichever property is picked, the point stands that (4) rests on the idea that Xia has *some* property essentially, and so presupposes the essentialist view that the distinction between essence and accident is in good standing.[10]

The opposing view, anti-essentialism, is that there is no intelligible distinction between essence and accident. The anti-essentialist does not say that all properties are had accidentally; that is something that only an (extreme kind of) essentialist could say. Rather, the anti-essentialist rejects all talk of essence and accident in the first place. Quine was an anti-essentialist. He famously rejected the intelligibility of *de re* modality, so he would reject the notion of essence defined above as unintelligible on the grounds that it "quantifies in" to modal contexts. Thus, Quine would reject claim (4) as unintelligible. He would accept the biological fact that if the couple waited then the resulting child would likely have different genetics, different origins, etc. But as an anti-essentialist, he would reject as unintelligible the claim that the child *would thereby not be Xia*.[11]

---

[9] I stress that this is just a working definition. Fine (1994) argued that this is not the correct definition, but those complications are not relevant to our purses so I bracket them here.

[10] This difference between saying that it is *impossible* for the child to exist had her parents waited, and saying that it is (merely) *unlikely*, is often ignored in the literature. Notable exceptions include Roberts (2007, 2009) and Hare (2013), who argue that the distinction is crucial. I am inclined to agree with them. More fully, if the problem is that Xia is *merely unlikely* to exist if her parents had waited, then, since her existence was *also* unlikely given when her parents chose to conceive immediately, they argue that we can deny (3) and say that their choice was bad for Xia because, at the time, her expected utility conditional on their conceiving immediately was less than it was conditional on their waiting. If that solution is workable, the hard problem arises if Xia's essence implies that it is *impossible* that she exist if they wait (and indeed in that case Roberts denies (1)).

[11] In the framework of possible worlds, essentialism is the view that there are facts of transworld identity. An essential property of Xia is then any property that Xia has in all worlds in which something is identical to her (where this could be a highly disjunctive or otherwise complex property).

So, (4) is not just a scientific claim about the biology of human reproduction. Rather, it presupposes the metaphysical thesis of essentialism. But essentialism comes in many varieties, and the view I want to develop becomes most visible when one looks at (4) through the lens of one particular variety. Let us zero in on that variety.

3. Ontic vs. Descriptive Essentialism

To this end, start by distinguishing *ontic* from *descriptive* essentialism. Suppose that an object x has property P. According to *descriptive essentialism*, whether x has P essentially or accidentally depends on the manner in which x is presented or described. By contrast, *ontic essentialism* is the view that whether x has P essentially or accidentally is independent of the manner in which it is described.

Given our definition of essence in terms of *de re* necessity above, this distinction amounts to a distinction between two ways of interpreting the latter. Thus, one descriptive essentialist view is that, when x is described with the description D, the right-hand-side of that definition is true iff:

necessarily, for all y, y = the D only if y has P.

Thus, if Obama is described as the 44th US president, then on this view he is essentially a president, since it is necessary that the 44th US president is a president. But when Obama is described as the father of Sasha and Malia, then on this view he is accidentally a president, since it is not necessary that the father of Sasha and Malia is a president. Thus, whether he is essentially a president depends, on this view, on how he is described.

Lewis's (1986) counterpart theory is another variety of descriptive essentialism. On this view, the right-hand-side of our definition is true iff every counterpart of x has P. And y is a counterpart of x iff y resembles x in those respects made salient by the context of conversation. So, in a context in which Obama's being a president is particularly salient, all of Obama's counterparts may be presidents, in which case Obama is essentially a president. But in a context in which his being a president is not salient, some of Obama's counterparts will be non-presidents, in which case Obama is accidentally a president. So, again, whether he is essentially a president depends on the manner in which he is being discussed.

By contrast, ontic essentialism is the view that whether x has P essentially or accidentally is independent of how x is described or presented. For our purposes, the key feature of ontic essentialism is that it implies that essentialist claims are *referentially transparent,* in the sense that all occurrences of the following inference form are truth-preserving:

1 x is essentially F
2 y is accidentally F

---

3 $x \neq y$

According to ontic essentialism, this form of inference cannot fail to preserve truth: if x is essentially F and $x = y$, then (by the indiscernibility of identicals) it must be that y is essentially F too. But according to descriptive essentialism, it can lead from truth to

falsity. For even if $x = y$, it may be that x is presented in one way in premise 1, and in a different way in premise 2, in such a way that both premises are true.

Which of ontic or descriptive essentialism is correct? I will not try to settle this issue in modal metaphysics. But the view I want to develop becomes most apparent on the ontic approach. So, for the time being, let us make the following assumption:

> Assumption A: If essentialism is true, then ontic essentialism is true.

I will later discharge this assumption, but it helps to simplify matters for now. Unless otherwise stated, I will use "essentialism" to refer to ontic essentialism.

4. Coincidence

Now, (ontic) essentialism leads to the view that there can be distinct yet spatially coincident objects. Think of a statue of Goliath, fashioned out of a lump of clay. Call the statue "Statue", and the lump "Lump". Consider Statue's statuesque shape. Does it have this shape essentially or accidentally? Arguably, essentially: if nothing had been fashioned into (roughly) this statuesque shape, then presumably there would have been no such thing as Statue. But now consider Lump, the lump of clay. It also has this statuesque shape. Does it have the shape essentially or accidentally? This time, we think accidentally: Lump could have been left alone and not molded into this statuesque shape.

But this means that the statue and the lump of clay are two distinct entities. After all, we just agreed that

(a) Statue is essentially statue shaped.

and that

(b) Lump is accidentally statue shaped.

By referential transparency, it follows that

(c) Statue $\neq$ Lump.

The statue and the lump of clay it is made out of are distinct entities, even though they spatially coincide. Of course, the *descriptive* essentialist denies referential transparency, so she may reject this argument that they are distinct. But, *ontic* essentialism leads to this conclusion that there can be distinct yet coincident entities.

To be clear, the ontic essentialist could resist this conclusion by rejecting (a) or (b). Thus, she could be a mereological nihilist and reject both (a) and (b) on the grounds that neither Statue nor Lump exist. Or she could join Burke (1994) and reject (b) on the grounds that Lump goes out of existence when fashioned into a statue. Alternatively, she could accept that both Lump and Statue exist but say, counterintuitively, that Lump is essentially statuesque, or that Statue is accidentally statuesque. But all these views give up intuitive and commonsensical beliefs of the form "there are Ks, and they are essentially F". And if you give up those beliefs, it is hard to see what confidence you could have in claim (4) of the nonidentity problem. After all, we believe (4) insofar as we believe that there are people, and that they have some property (such as their origins, or genetic code, or date of conception) essentially. Thus, in the current context, in which

we are supposing that (4) is true and examining its metaphysical underpinnings, it would be dialectically odd to reject (a) or (b). So the point is this: ontic essentialism, plus the kinds of background beliefs needed to generate (4) in the first place, yield the view that there can be distinct yet coincident entities.

Note that this is not the spooky view that statues and lumps are like ghosts. We have not just shown that two objects *composed of different matter* can spatially coincide—that would indeed be spooky! Rather, the view is that the objects of ordinary thought and talk are not mere parcels of matter, but are parcels of matter *together with* a specification of their essential properties. This is the modern reincarnation of the Aristotelian idea that an ordinary object is a composite of matter and form. Seen like this, it is unsurprising that distinct objects can share the same matter, but differ with respect to which properties they have essentially (their "form").

5. Unlimited Essentialism

The next question is how far this multiplication of coincident objects goes.

To see the issue, suppose that Statue is in my office. Could it have been elsewhere? Presumably it could—we normally think that statues are the kind of things whose locations are accidental, so that they can be moved around without destroying them. But is there *another* coincident object that is exactly like the statue with the *one* difference that it is in my office essentially? Admittedly, this object is a little strange: if you took it out of my office, you would destroy it—though you would not destroy Statue or Lump! So perhaps it is not the kind of object we ordinarily pay attention to. But it is not unimaginable that there be such a thing—indeed, works of installation art arguably have their locations essentially. So the question might be put like this: in fashioning a statue out of some clay, could one create *two* works of art in one go, an ordinary statue and an installation piece?

The question reiterates. Take the set of all (non-modal) properties of Statue, including its shape, color, location, weight, and so on. Statue and Lump share those very same properties; they differ only with regards to which subset of these properties they have essentially. Then the question is this: Does *any* subset of these properties correspond to the properties that are had essentially by *some* entity coincident with Statue?

According to *limited essentialism*, the answer is "no". There is a statue, a lump, and perhaps an installation art piece. But that is all: most subsets of the statue's non-modal properties are not had essentially by anything. By contrast, *unlimited essentialism* answers "yes". On this this view, there is a dazzlingly large number of distinct entities coincident with Statue, one for each subset of its non-modal properties! There is an entity that is essentially in my office, which is destroyed when I take it elsewhere; an entity with its value essentially, which is destroyed when its value rises; an entity with its position relative to Sagittarius essentially, which is destroyed as we rotate around the sun; and so on. According to unlimited essentialism, the same goes for all ordinary objects, including people. In addition to Obama, sitting in the Oval office, there is also sitting-Obama, an entity that has exactly Obama's non-modal properties and differs only in the fact that it has the property of sitting essentially, rather than accidentally. When Obama stands up, sitting-Obama goes out of existence! To foreshadow, there is also an entity that has exactly Obama's non-modal properties and differs only in the fact that it has its origins (and genetic code, and date of

conception, etc.) accidentally, not essentially. According to unlimited essentialism, all these entities are out there.

Is limited or unlimited essentialism correct? One might complain that unlimited essentialism posits more entities than are recognized by commonsense. But the unlimited essentialist has a ready explanation of this fact: for pragmatic reasons, everyday thought focuses just on those few entities that are useful to represent when navigating the world. Against limited essentialism, the question is whether a principled distinction can be drawn between those sets of properties that constitute the essential properties of an entity, and those that do not. Many contemporary essentialists—including Kit Fine (1999), Mark Johnston (2006), and Kathrine Koslicki (2008)—are skeptical that a principled distinction can be drawn, and are unlimited essentialists as a result.[12]

I will not settle the issue here. My aim, remember, is to isolate that variety of essentialism on which the view I want to develop becomes most visible, and that variety is unlimited essentialism. So let us make a second assumption:

Assumption B: If ontic essentialism is true, then unlimited essentialism is true.

Like the first assumption, I will discharge this later on.

6. The Breakfast People

Go back to our nonidentity problem. The puzzle was that these four claims are individually plausible but jointly inconsistent:

(1) Conceiving immediately was morally wrong.
(2) Conceiving immediately was morally wrong only if it was bad for Xia.
(3) Conceiving immediately was bad for Xia only if Xia would have existed had they waited.
(4) Xia would not have existed had they waited.

We are now in a position to describe the view I want to develop. The idea is that, if Assumptions A and B are granted, then coincident with Xia are a *multitude* of entities differing only in which properties they have essentially. According to the view I have in mind, some of these entities make (4) true, others make (2) and (3) true, but none make them all true together.

To see how this works, consider a fictional community called the Breakfasters. The Breakfasters are much like us; they differ only insofar as they have a different conception of people. They agree with us that something like a person's origin, or genetic code, or date of conception, are essential properties of them. So, they agree with (4) that if the couple had waited a month before conceiving, they would have had a different child.

[12] See Bennett (2004) for a discussion and application of unlimited essentialism. There are a number of subtle questions about to properly formulate unlimited essentialism. For one thing, an unlimited essentialist may want to say that if the property of being red is had essentially by x, then the property of being colored is also had essentially by x. In that case, the view cannot be that *any* subset of its non-modal properties constitutes the essence of *some* distinct entity—some "consistency" constraint must be imposed. See Yablo (1987) for more on this issue. But for our purposes there is no need to refine the view precisely here; it is enough that we have the picture.

The difference is that they think that what someone's biological mother ate for breakfast on the 10th day of pregnancy is *also* one of their essential properties. Suppose my mother ate cereal for breakfast on the 10th day of pregnancy with me. Then the Breakfasters think that if she had eaten eggs, she would have had a different child; she would not have had *me*. So, the Breakfasters think that choosing what to eat for that meal is like choosing when to conceive: it is (in part) choosing which child to have.

Imagine, then, a Breakfaster named Sabita who is 10 days pregnant. She has two breakfast options: cereal, or cheese on toast. She likes both but mildly prefers cheese on toast. But the cheese is a particularly potent strain that contains a substance known to affect a 10 day-old fetus. If she eats the cheese, her child will be born with the very same condition as Xia, a condition that produces a poor quality of life. Sabita knows all this, and reasons as follows:

> "The situation is puzzling. On the one hand, it would seem wrong for me to eat the cheese. But it's wrong to eat the cheese only if it's bad for my child. And, on reflection, it *won't* be bad for my child. True, my child will live a hard life, one that is barely worth living. But if I ate the cereal instead, I would have a different child. Thus, far from being bad for the child, my eating the cheese will be a necessary condition for her existence! So, on reflection, it's not wrong to eat the cheese after all."

Sabita faces a formally analogous "nonidentity problem". Suppose she goes ahead and eats the cheese, and gives birth to a child with the condition. Call the child "Tia". Then her nonidentity problem is that the following four claims are each plausible to her, yet inconsistent:

(1*) Eating the cheese was morally wrong.
(2*) Eating the cheese was morally wrong only if it was bad for Tia.
(3*) Eating the cheese was bad for Tia only if Tia would have existed had Sabita eaten cereal instead.
(4*) Tia would not have existed if Sabita had eaten cereal instead.

Given Sabita's strange views about essence, she believes (4*). And her speech shows that she also believes (2*) and (3*). So her conclusion is that it is not wrong to eat the cheese after all, *contra* (1*).

What should we say about Sabita's reasoning? Clearly, she made a mistake. For one thing, many would say that her conclusion is false, insisting that it is in fact wrong to eat the cheese. But even if you agree with Sabita that it is OK to eat the cheese, it seems bizarre to establish that conclusion on the basis of her strange essentialist beliefs encoded in (4*). If it is OK to eat the cheese, that is presumably because of considerations pertaining to Sabita's rights as a mother or something of that ilk. Thus, regardless of whether her conclusion is true or false, it seems clearly a mistake to draw that conclusion *on the basis of her strange essentialist beliefs encoded in (4*).*

The same goes if Sabita had rejected (2*) or (3*) on the basis of (4*). There may be good reasons to reject (2*) and (3*). But it is a mistake to reject them *on the basis of her bizarre essentialist beliefs encoded in (4*).*

But why, exactly, is that a mistake? Is it because (4*) is false, so that she is arguing from a false premise? No—when Sabita asserts (4*), she speaks truly! Remember, we

are assuming unlimited essentialism, so there are indeed entities out there that her strange beliefs about essence are true of. These entities coincide with people. Indeed they are just like people, except they are entities for whom what their mother had for breakfast on the 10th day is an *essential* property of them, not an accidental property. Call them schmeople. We do not ordinarily think or talk about schmeople: names in *our* language denote entities for whom what their mother had for breakfast on the 10th day is an *accidental* property. But the Breakfasters think and talk about schmeople, not people, so when they introduce proper names like "Tia", they to refer to schmeople.[13] So, when Sabita asserts (4*), she speaks truly.

Sabita's mistake, then, is not that (4*) is false. Instead, I claim that her mistake is to assume that schmeople are the appropriate things to be concerned about when deciding what to do. What do I mean? Consider the Person Affecting Principle that supports (2*), on which an act is wrong only if it is bad for *someone*. What kind of entity does "someone" quantify over? According to unlimited essentialism, there are indefinitely many candidates out there: people, schmeople, and so on. Which of these must a wrong act be bad for? Sabita's mistake was to assume that a wrong act must be bad for some *schmerson*, so that when thinking about whether an act is wrong we must consider how it affects schmersons. That is clearly a mistake: schmersons are not what matter, they are not morally significant entities. If (2*) is true, it is true of some other kind of entity.

The same goes for (3*), which concerns the conditions under which acts are bad for people. What exactly are the conditions at issue here? What kind of entity can an act be *bad for* in the morally relevant sense? According to unlimited essentialism, there are many candidates out there: people, schmeople, and so on. Sabita's mistake was to assume that if acts are bad for anything, they are bad for schmeople. Again, that is clearly a mistake: schmeople are not what matter. If (3*) is true, it is true of some other kind of entity.

Thus, the thing to say about Sabita's nonidentity "problem" is this. If she ate cereal instead of cheese, she would indeed give birth to a different schmerson. So, (4*) is true of schmersons. But schmersons do not matter. The entity that matters, in Sabita's situation, does *not* have as an essential property what its mother had for breakfast on the 10th day. Who knows exactly what this entity is—perhaps it is a person, perhaps not. But whatever it is, *it* (the self-same thing) would be born no matter what Sabita had for breakfast. *If* (2*) and (3*) are true, they are true of *this* entity. So interpreted, (1*)–(4*) are not inconsistent. Her nonidentity "problem" has dissolved.

7. Flexistentialism

According to the view I want to develop, we have been making the same mistake as Sabita. Consider Xia, the child that our couple had after conceiving immediately. There are indeed entities coincident with Xia that make (4) true. These are entities that would not exist had the couple waited a month, because they have (something like) their origins, or genetics, or time of conception essentially. In this respect they resemble workaday, three-dimensional, medium-sized dry goods like tables, chairs, pens, and flowers, which are also thought to have something like their circumstances of origination

[13] If this is not clear, just embellish the case. Suppose that the Breakfasters introduced their term "person" explicitly to refer to entities that have as essential properties what their mother had for breakfast on the 10th day; and then suppose that when they name someone they say "Let this name denote that *person*".

essentially. Thus, let us call the entities that make (4) true *workaday person-like entities*, or "person-likes" for short.[14] For whatever evolutionary or social or political reasons, we often focus our attention on person-likes, which explains why we so naturally think that (4) is true.

But according to the view I want to develop, person-likes do not matter. They are no more morally significant than entities that have as essential properties what their mother ate for breakfast on the 10th day of pregnancy. Which entities matter? On the view I have in mind, what matters in our nonidentity problem is an entity coincident with the person-likes but with fewer essential properties, perhaps just the property of *being their first child*. This entity does *not* have its origins (or genetics, or date of conception, or anything of that ilk) essentially, so that *it*—the self-same thing—would be born no matter whether the couple conceives immediately or waits. *If* (2) and (3) are true, they are true of *this kind of entity*: it is wrong for the couple to conceive immediately because it is bad for *it*, in the straightforward sense that *it* would have been better off had they waited. So interpreted, (1)–(4) are not inconsistent: (4) is true of the person-likes, but not the entity that matters; (2) and (3) are true (if true at all) of the thing that matters, but not the person-likes. To be clear, it may be that (2) and (3) are false even when interpreted as being about the entity that matters. The point is that on this view, *if* they are true that would not be inconsistent with (4); our nonidentity "problem" is dissolved.

Thus, the view is that what makes (4) true is one thing, and what matters is a *different* thing, and one finds the latter by adopting a flexible conception of which properties are had essentially by the things that matter. So the view could be called "flexible essentialism", or perhaps "flexessentialism" for short. But even the latter is a little clumsy. Thankfully, since essence fixes existence-conditions, it should not be too misleading to call it *flexistentialism* instead.

When I say that an entity *matters*, I mean that it is an appropriate object of practical concern. Practical concern includes moral concern, so an entity matters if it is the kind of entity one should consider when thinking *morally* about what to do—when thinking how one's actions will affect others in morally relevant ways. But practical concern also includes non-moral concern, so something matters if it is the kind of entity one should think about when thinking *self-interestedly* about what to do—when thinking about what is best *for one's self*. *Mutatis mutandis* for other kinds of non-moral practical reasoning. I will sometimes use the term "practical person" as a synonym for "entity that matters". It is then an open question which entities matter; that is, what their essential properties are. We overlook this question when we assume that, faced with someone, there is only one entity there. For we can all agree that *something* there has its origins essentially, so that if there is only one thing, *that thing* must matter. Case closed. But if unlimited essentialism is true, this is to overlook a foundational question of ethical theory: namely, which of the multitude of coincident entities matter.

Flexistentialism offers a partial answer to this question. It says that, in our nonidentity case, the entity that matters is not a workaday person-like entity: it does not have as an

[14] This is a riff on Mark Johnston's term "personites" (Johnston, manuscript), which he uses to refer to short-lived entities that spatially coincide with a person at a given time, and which persist through part but not all of that person's life. In contrast to a personite, a person-like may persist through the whole of a person's life.

essential property anything that would make (4) true. This does not imply that we should ignore the person-likes altogether. Insofar as person-likes resemble artifacts, animals, and other material things in their essential profile, thinking about person-likes may be important when locating ourselves within the material world—a project that Velleman (2008, pp. 262) speculates is important. Relatedly, the flexistentialist can agree that there are many useful concepts of person: perhaps one relevant to identity politics, when we say (for example) that to loose one's religion is to undergo a change in one's identity; perhaps yet another concept used on January 1st when we say "I resolve to be a new person this year!" All these may be legitimate concepts of "person". Flexistentialism is just the idea that these concepts may come apart from the concept of a *practical* person, i.e. the concept of an entity that matters in the sense outlined above.

So far, the flexistentialist has told us what the practical person in our nonidentity problem is *not*. But she has not said much about what it *is*, except that it is something that would exist had her parents waited. There are a number of possible views about what this entity could be, and I will not try to choose between them here. But it is worth discussing some options.

First, we should distinguish universalist from particularist approaches. On the *universalist* approach, all practical persons have the same kinds of essential properties. Indeed, this universalist approach is presupposed in the literature on the non-identity problem, where it is assumed that the entities that matter all have their origins essentially, or that they all have their genetics essentially, and so on. By contrast, a *particularist* approach allows that the essential properties of practical people may differ from case to case. Thus, the particularist might say that in our nonidentity case, the thing that matters just has the property of *being their first child* essentially; but that the thing that matters in other cases may have some other kind of essential property.

One might think of particularism in the following way. Parents stand in a special relationship with their children. And one might think that it is because of that relationship that parents have particular responsibilities to their children, and conversely that their children have certain reasonable expectations of their parents. So a particularist might say that this relationship defines the entity that matters. More generally, the view would be that the relationships that ground our moral commitments determine the essential properties of the entities that matter. Since those relationships vary from case to case, so do the kinds of entities that matter.[15]

I will assume particularism here, not because I am certain of its truth but because it is a convenient working hypothesis. The question, then, is this: In our nonidentity case, which of the many entities coincident with the couple's child matter? It may be that many of them matter, but let us assume for simplicity that exactly one entity matters. And from here on, let us stipulate that "Xia" is to refer to this entity—until now the name has been ambiguous between the person-likes and the entities that matter. Then the question is: Which entity is Xia? It is obviously not enough to say "The person over there", pointing at the couple's child, since there are many coincident entities, and it is unclear which one "person" picks out. What, then, must be specified in order to uniquely pick out Xia?

At a minimum, one must specify her essential properties. So, we might say that Xia is the one whose only essential property is *being their first child*. But this is not enough.

[15] I am indebted to Louis-Philippe Hodgson for showing me that this kind of particularist view is possible.

To say that this is an essential property of Xia is just to say that having this property is a *necessary* condition for being Xia. But is it also a *sufficient* condition for being Xia? On one view, it is: in any world that the couple have a first child, that child is Xia. But on another view, *being their first child* is a necessary but not sufficient condition for being Xia: there is a world in which the couple have a first child, but it is not Xia. To be clear, both entities may be present and coincident: an entity x such that *being their first child* is necessary and sufficient for being x, and a distinct entity y such that *being their first child* is necessary but not sufficient for being y. A statement of what Xia is (or what "Xia" refers to) must say which of these entities she is.

Thus, say that a property P is an individuating property of x iff having P is sufficient for being x. That is:

P is an *individuating property* of x iff necessarily, for any y, if y has P then $y = x$.

Then to specify what entity matters, one must list all its essential properties *and* its individuating properties. Call these two lists its *essential profile*. Then one flexistentialist view is that Xia is the entity for whom *being their first child* is its one and only essential property, and its one and only individuating property. As desired, this is an entity that would still have been born had the couple waited.[16]

This view is convenient to have in mind for the sake of concreteness, but I stress that it is only one amongst many flexistentialist views about our nonidenity problem. What all flexistentialist views have in common is that they say that Xia (the entity that matters) has an essential profile on which she would be born even if her parents had waited a month. Of course, now that we have stipulated that "Xia" refers to the entity that matters, this implies that (4) is false. But it is misleading to describe flexistentialism as the view that (4), ordinarily understood, is false. For we originally introduced the name "Xia" in section 1 with the reference-fixing description "their child", and it is not at all clear whether *that* stipulation picked out the entity that matters. Whether it did depends on what kinds of entities satisfy the English predicate "child"—person-likes, or practical persons, or both—and I will not discuss that semantic question here. So, put aside the semantic question of whether (4) is true in ordinary English: the flexistentialist's substantive claim is that, while there are entities out there that make (4) true, they are not the entities that matter.

No doubt flexistentialism sounds odd at first; absurd, even. It implies that if the couple had waited, Xia would have been born a month later with very different genetics. This is runs counter to what we ordinarily think about people, to put it mildly! But is that a problem? Remember, Sabita would find it just as odd to say that she would have had the same child had she eaten a different breakfast, but we know this is what she should say (insofar as she is talking about the entity that matters). Moreover, flexistentialism

[16] I am bracketing subtle issues about how to individuate properties. Suppose one thinks that co-intensive properties are identical. Then, if *being their first child* is essential and individuative of Xia, it follows that the property of being their first child = the property of being Xia. Something must then be said about why it is informative to specify the practical person in one way but not another. By contrast, if one thinks that co-intensive properties can be distinct, then Xia will have many essential properties indeed: the property of being their first child, the property of being their first child and being such that 2+2=4, and so on. Something must then be said about how we can effectively communicate essential profiles in a finite manner. But a full discussion of these issues would distract us from the main thread.

acknowledges that there are entities coincident with Xia (the person-likes) that would *not* have been born if her parents had waited; insofar as we normally attend to those entities, this explains why we have the ordinary beliefs we do. In any case, I will discuss reasons to accept or reject flexistentialism later. For now, put aside whether you think flexistentialism is *true*: my claim so far is just that it is a *coherent* position in logical space.

Still, its coherence is enough to show that the standard approach to the nonidentity problem is based on a mistake. The standard approach assumes that (1)–(4) are jointly inconsistent, so that accepting (4) *commits* one to rejecting one of (1)–(3). This is a mistake because it overlooks the possibility that (4) is true, though not of the entities that matter; so that (1)–(3) may *also* be true, when understood as concerning the things that matter. More generally, what this shows is that there is no direct route from this nonidentity case to a rejection of the Person Affecting Principle, or the Counterfactual Comparative view of "bad for". For one can accept that both principles are true of the entities that matter, while also recognizing the case as a "nonidentity case" insofar as the couple's choice affects the identity of the workaday person-like entities.

8. Other Cases

That completes my primary aim of showing that flexistentialism is a coherent option. My secondary aim is to show that there are good reasons in favor of flexistentialism, so that it might actually be true. But before that, let us zoom out briefly. We know that Xia's nonidentity problem is one of many. What happens if we apply the flexistentialist approach to other nonidentity problems? Let us examine three representative cases.

Consider first a case in which a woman—call her Mary—is choosing whether to have a child immediately with the same condition as Xia, or wait 10 years and have a child without the condition with a different man. In this case, the flexistentialist view is that the entity that matters is one that would be born regardless of Mary's decision; for example, that it is an entity for whom *being Mary's first child* is its essential and individuating property. Call this entity Yena. Note that Yena could have had a different biological father! This is an odd thing to think about someone, for sure; but as before let us wait until we consider arguments for and against flexistentialism before casting judgement.

The second case involves Mike, who works in a lab for a company that offers *in vitro* fertilization. He has a sperm and an egg from a couple in Spain, and a sperm and an egg from a couple in China. But Mike can only fertilize once, so he must choose between the couples. There is no reason to favor one couple over the other, with one exception: the Spanish couple's gametes will produce a child with the same condition as Xia, but the Chinese couple's gametes will not.[17] Ordinarily, we would think of this as a nonidentity problem, in which the identity of the resulting child depends on Mike's choice. But the flexistentialist solution is to say that the entity that matters is an entity that would be born no matter which couple Mike chooses. For example, the flexistentialist might say that it is an entity for whom *growing from the zygote formed by Mike* is its essential and individuating property. Call this entity Zeta. Suppose Zeta is in fact born to Chinese parents and grows up in Chengdu. Then the flexistentialist insists that she *could* have been

[17] Thanks to Caspar Hare for describing this kind of case to me.

born to different parents in Madrid! It might sound incredulous that Zeta really could have lived such a wildly different life, with different biological parents. But again, put that aside until we consider arguments for and against.

Finally, it will be instructive to consider the more complex case of environmental conservation described in the introduction. Suppose we do not conserve, and in 300 years our descendants are eeking out a poor existence on our ravaged planet. Did we do wrong? Intuitively so, but the puzzle was generated by the thought that if we had conserved, those descendants would not have existed. The flexistentialist solution would agree that the *workaday person-like entities* would not have existed, but that the entities that matter are different and would exist regardless of what we did. But what could these entities be? Consider one of our descendants. Call her "Wanda", on the understanding that this denotes the entity that matters. What could Wanda's essential profile be? We know that if we had conserved, the resulting people (whoever they are) would live lives that are *unimaginably* different from Wanda's actual life, enjoying luxuries and pursuing careers that she will never dream of. They would have very little in common with the actual Wanda, other than being a future human. So, for Wanda to *be* one of those resulting people, she must have almost all her properties accidentally. Perhaps her only essential property is *being a future human*.

What about her individuating properties? What would make one of those counterfactual individuals *Wanda*? There are many possible views; let me describe just two. One is a "Cartesian" view, on which Wanda's only individuating property is the unanalyzable, irreducible property of *being Wanda*. On this view, an exact qualitative description of how things are in 300 years time leaves open which future individual is Wanda; it is a further, brute fact who Wanda is. This allows that it is *possible* that Wanda exists if we conserve. But it is consistent with this view that many worlds in which we conserve do not contain Wanda, so that if we had conserved, Wanda would *likely* not exist. Thus, whether this view helps solve the nonidentity problem depends on what the problem was. As noted in section 2, there are two kinds of nonidentity problem: one based on the idea that it is *impossible* for Wanda to exist if we conserve, the other based on the idea that she would *likely* not exist if we had conserved. The Cartesian view helps solve the first problem, but not the second.

A second view, which might be called the "reductive role" view, does better on this count. On this view, Wanda's individuating properties are qualitative roles. Consider a world W in which we conserved. And consider a complete, qualitative description of humanity in W, stating that there is one person playing role R1, another playing role R2, and so on. These roles might be very complex, but one might nonetheless say that one of them is an individuating property of Wanda, so that, necessarily, anything playing this role is Wanda. On the reductive role view, Wanda's complete list of individuating properties is then a list of such roles, one for each world in which she exists. As long as the list includes enough roles, this will imply that she exists in most, or even all, worlds in which we conserve, so that—unlike on the Cartesian view—she would (likely) exist if we had conserved.

Now, the case of environmental conservation differs from the rest in being a "different number" case: a case in which the *number* of resulting people likely depends on our choice. This raises a worry. Suppose that, had we conserved, fewer people would have existed in 300 years. Then how can it be true, of each actual descendent, that they would have existed if we had conserved? The relevant counterfactual worlds do not contain

enough bodies to go around! The solution is to remember that we do not need as many counterfactual bodies as there are actual descendants. The reductive role theorist can allow that Wanda's list of individuating properties overlaps with her friend Violet's list, so that in some counterfactual worlds they both exist in the same body. This might sound absurd, but is a natural consequence of the unlimited essentialist view that there are many entities coincident with any material body, differing only in their modal profiles. The resulting view raises many questions, but if it can be sustained it shows that the reductive role view can accommodate the idea that all our actual descendants would exist, even if we had acted in such a way that led to fewer live human *bodies* in 300 years time.

This is just a sketch and leaves out many details. But I hope to have indicated what kinds of views we end up with if we apply flexistentialism to these other nonidentity cases. The views are wild, no doubt, but it seems to me that they are *coherent*.

Let us now go further and ask whether they might be *true*.

9. Three Easy Pieces

Flexistentialism might be true of some nonidentity cases but not others, and I will not argue here that it is true of *all* cases. So let us focus again on our original case of the couple that conceived immediately and had Xia. It should be clear how the arguments pertaining to their case generalize, but I will not discuss whether they are equally compelling in the other cases. And let us continue to make Assumptions A and B, so that coincident with Xia are many entities differing in their essential profiles; I will relax these assumptions later on. Our question is which of these many entities is *Xia*, the entity that matters. Flexistentialism says that she is not a workaday person-like entity, but an entity that would have existed regardless of whether the couple waited or conceived immediately. I will outline four arguments for this view. None is decisive, and I do not accept them all myself, but I present them to set out the flexistentialist's options. Together, they suggest that flexistentialism is an approach to take seriously.

The first argument proceeds on the basis of the moral principles articulated in our initial presentation of the nonidentity problem. Suppose you believe, with (1), that conceiving immediately was morally wrong. And suppose you believe the Person Affecting Principle, and so agree with (2) that it was morally wrong only if it was bad for Xia (that is, bad for the entity that matters). Suppose, finally, that you believe the Counterfactual Comparative view of "bad for", and so agree with (3) that conceiving immediately was bad for Xia only if Xia would have existed if they had waited. It follows that Xia would have existed if they had waited, and is *not* a workaday person-like entity, just as flexistentialism states. Call this the "argument from moral wrong".

One might complain that this is disappointingly simplistic or *ad hoc*, a desperate move motivated only to avoid paradox. But that mistakes the dialectical situation. As we have seen, there is a genuine question—indeed a fundamental question of ethical theory—as to which entity matters. The literature on the nonidentity question simply *ignores* this question, and considers it an *obvious truism* that the entity that matters is one that would not have been born had her parents waited. Seen like this, it is (1)–(3) that constitute the real meat of the paradox, and I agree that denying (4) then appears desperate and *ad hoc*. But the dialectic is very different once we appreciate that there is a difficult and substantive

question as to which of the multitude of coincident entities matter; that is, the question of which entity (1)–(3) are *about*. Seen like this, (1)–(3) are not paradoxical *at all*, but rather function as premises in a straightforward argument that the entity that matters would be born no matter whether the couple waits or conceives immediately. There is nothing *ad hoc* or desperate about it at all.

Still, I doubt that the argument will convince many, since most theorists who have considered the matter reject (2) and hence reject the Person Affecting Principle, or reject (3) and hence reject the Counterfactual Comparative view of "bad for". Of course, if one rejected them because of the nonidentity problem, it would be question-begging to reject the argument from moral wrong on that basis! But still, some theorists reject (2) and (3), and the associated principles, for other reasons, and perhaps they are right to do so. So let us find other arguments that do not appeal to these claims.

It is easy to find an argument that does without (2) and the Person Affecting Principle. For regardless of (2), it remains natural to think *that conceiving immediately was bad for Xia*. This is a natural thought even if one denies that this was a necessary condition on their action being morally wrong (indeed, even if one denies that their action was morally wrong). And this italicized claim is enough on its own to drive a second argument, an "argument from badness". For, along with (3), it follows that Xia would exist if the couple had waited, just as flexistentialism states. Of course, some defense of the italicized claim should be offered here, but I will leave that to flexistentialists who wish to take this argument up.

I will present an argument that does without (3) in due course. But before that, note that the two arguments just described—from moral wrong, and from badness—turn the standard approach to the nonidentity problem on its head. The standard approach takes the *metaphysical* thesis behind claim (4) for granted—that the entity that matters has, say, its origins essentially—and uses it to reject the *ethical* theses of (1), (2), or (3). This is a "metaphysics-first" approach: it treats our pre-theoretic, metaphysical opinions about the essential profiles of people as a sacrosanct core around which our normative theorizing should be based. In contrast, the two arguments above go in the reverse direction. Their premises are *ethical* claims about what is wrong, or about what is bad for whom, and they draw *metaphysical* conclusions about the essential profiles of the entities that matter. If a given hypothesis about Xia's essential profile makes best sense of our ethical beliefs, then this approach accepts the hypothesis even if it runs against pre-theoretical opinion. On this "ethics-first" approach, the metaphysics of persons cannot be done in isolation of ethics. As the saying could be put now, there is no such thing as ethics-free metaphysics; there is only metaphysics whose ethical baggage has been taken on without examination.[18]

It is surprising that this ethics-first approach has been largely ignored in the nonidentity literature. For it mimics a standard approach to the question of personal identity over time. There, theorists rarely *start off* with fixed assumptions about what it takes to survive. Rather, it is standard practice to use normative and evaluative judgments about what futures one should prefer, or about who is responsible for which past actions, to guide our theorizing about what it takes to survive. The ethics-first approach adopted here is analogous: rather than start off with fixed assumptions about what our essential

[18] With more apologies to Dan Dennett; see footnote 8 for the real quote.

profiles are, it uses normative and evaluative judgments to guide our theorizing about the essences of the things that matter.[19]

A third argument for flexistentialism—the "equivalence argument"—also follows this ethics-first approach. Suppose that if a pregnant woman shoots heroin in the third trimester, her child will be born with Xia's condition. If so, it is tempting to think that it would be morally wrong to shoot heroin in the third trimester. Indeed, Parfit claimed that if Xia's parents waited and then shot heroin in the third trimester, that would be *just as* morally wrong as conceived immediately, no more no less. This was his "No Difference" thesis. What explains this equivalence?

The flexistentialist's explanation is straightforward. In both cases, the relevant act (conceiving immediately, or waiting and then shooting heroin) is bad for the entity that matters in the Counterfactual Comparative sense that *that very entity* will be worse off than it would have been had the parents done otherwise. By hypothesis, the entity will be worse off to the same extent in both cases; hence the two acts are equally wrong.

But on other views, the explanation is less straightforward. To see this, let us make the plausible assumption that shooting heroin is wrong in part because it is bad for the child in the Counterfactual Comparative sense. The trouble is that a non-flexistentialist cannot say the same about conceiving immediately. Thus, she must (i) provide an alternative reason why conceiving immediately is morally wrong, and then (ii) argue that those reasons make it *just as* wrong as shooting heroin. And while there have been attempts at (i) by theorists who reject claims (2) or (3), I have not seen a sustained attempt to establish (ii). Moreover, even if one establishes (ii), the resulting explanation of the equivalence is a Rube Goldberg machine compared to the flexistentialist's. The flexistentialist can boast that her explanation is simpler and more unifying, and therefore better. That is the equivalence argument.

The point can be underscored by a comparison with Sabita. *We* do not draw any kind of moral distinction between eating cheese on the 10th day of pregnancy and shooting heroin in the third trimester (assuming that their consequences are the same). Sabita did, but only because she had mistaken beliefs about the essential properties of practical persons. Her mistaken beliefs got in the way of her seeing the moral equivalence between the situations. The equivalence argument asks us to consider that our parochial beliefs about essence are stopping us from seeing the moral equivalence between conceiving immediately and shooting heroin in the third trimester.

Of course, one might respond by rejecting the moral equivalence between conceiving immediately and shooting heroin. But if one rejects it because the former is a genuine nonidentity case and the other is not, one begs the question. We all agree that shooting heroin in the third trimester does not affect the identity of the entity that matters; the question at issue is whether conceiving immediately (rather than waiting a month) does. So it is question-begging to assume that it does, and then reject the moral equivalence *on that basis*. To assess the moral equivalence, we must put aside our theoretical beliefs

[19] This kind of "ethics first" approach is familiar in many other debates too. Consider debates about the ethics of marriage. A "metaphysics first" approach would first settle the question of *what marriage is*—what its essence or definition is—and from there derive ethical results about whether, say, marriage should be extended to gay couples. But the reverse "ethics first" approach is nowadays more standard: it first asks what legal solution is most ethical, and from there draws conclusions about how marriage must be defined to suit that legal solution.

about the identities of the resulting children. But I leave further defense of the equivalence to flexistentialists who wish to use this argument.

The argument also assumed that shooting heroin is bad in part because it made the child worse off in the Counterfactual Comparative sense. One might deny this, and so maintain that in *both* cases the wrongness of the act has nothing to do with making the child worse off in this sense. The resulting explanation of the equivalence may then rival the flexistentialist's in simplicity and unification.[20] Fine; but suppose that aggressively prodding the child moments after birth produces the same condition as Xia. We could run the equivalence argument with this case instead. Can one really maintain that the wrongness of prodding has *nothing to do* with making the child worse off? That strikes me as implausible, but again I will not defend this assumption in detail here.

10. Why Person-likes?

Before I present a fourth argument for flexistentialism, let me pause to ask why one might *reject* flexistentialism. Continue to make Assumptions A and B, so that there is this host of coincident entities. Why, in light of the arguments just presented, would one think that it is the *workaday person-like entities* that matter? Why, out of all of the myriad entities, would the things that matter have their *origins* essentially? *Mutatis mutandis* for any other essential properties of the person-likes, such as genetics, date of conception, and so on. What is so special about *those* features?

It is not at all clear. Admittedly, we have already seen how counterintuitive flexistentialism is, especially in the other nonidentity cases described in section 8. We are pre-theoretically disposed to think that Xia (the thing that matters) would not have been born had her parents waited, that Wanda would not exist had we conserved the environment, and so on. But are these pre-theoretic opinions epistemically significant? I suggested not, for two reasons that I should clarify. First, the case of Sabita functions as a skeptical scenario: if she unwittingly made a mistake, why think that we do better? And second, our pre-theoretic opinions are easily explicable. After all, we believe (i) that *something* in the vicinity of a person is a person-like, i.e. has its origins, or genetics, or date of conception, essentially; and (ii) that *something* there matters. These beliefs are correct. Put them together with the belief (iii) that there is only *one* entity in the vicinity of a person, and we get our pre-theoretic opinion that the thing that matters is a workaday person-like. Thus, our pre-theoretic opinion is easily explained. The problem is that (on current assumptions) it rests on a false presupposition, namely (iii). The opinion should therefore count for little, if anything. We need arguments.[21]

Let us suppose that one's *body* is the kind of thing that has its origins essentially, just like tables and chairs. If one then thinks that one is identical to one's body, we get the view that practical persons are workaday person-likes. But why think that you are identical to your body? Many think one can survive a "body transplant", in which one's body is (perhaps gradually) replaced with another. This is implied by the popular "psychological

---

[20] This was Parfit's (1986) own explanation of the equivalence.

[21] One might object that pre-theoretic opinion also includes the direct and basic insight that the things that matter are workaday person-likes. Put otherwise, the idea would be that when we pre-theoretically think that conceiving immediately is bad for Xia, we are thinking that if it is bad for *anything*, that thing must be a workaday person-like. I find the suggestion extremely implausible: I very much doubt our pre-theoretic intuitions are so finely grained.

continuity" view of personal identity over time, on which one survives an event just so long as one's pre-event psychological states are connect to the post-event psychological states in the right way. This view clearly implies that you are not identical to your body. Indeed, if one is accepts the psychological continuity view, it would be very bizarre to think that you are something that has its bodily origins essentially! There is thus something of a tension with thinking, with Parfit (1986), that it is psychological continuity that makes for your survival over time, yet you have your bodily origins essentially.

Moreover, even if you accept the "body" view of personal identity over time, on which one survives an event just so long as one's body does, it does not follow that you are identical to your body or that you have your bodily origins essentially. All that follows is that, whichever body you happen to be born with, you cannot survive the loss of that body. But that is consistent with the view that it is possible for you to be born with a different body; that your bodily origins are not essential to you.

Thus, the literature on personal identity over time does not obviously suggest a reason to think that we are workaday person-likes. One might hope for more from the literature on "moral standing". There is a clear difference between something that deserves ethical treatment—a person, a dog perhaps—and something like a pebble that does not. The literature on moral standing asks what makes an entity fall into the former category, and this is exactly the question we are asking.

But on closer examination the literature on this question is of little help to us here, because it presupposes that there is no explosion of coincident entities. For example, one view of the matter is that x has moral standing iff x is sentient. If there is no explosion of coincident entities, this may count as an answer to the question of moral standing. But if there are coincident entities, it is a partial answer at best. For coincident with any sentient being will be a huge number of other entities, differing only in their essential profiles, that are also sentient. A full answer to the question of moral standing—and an answer to our question of which entities matter—must say which of these coincident entities have moral standing, and the literature gives little guidance here.

Thus, if there is an explosion of coincident entities, it is unclear why one would think that the ones that matter are workaday person-likes. Put that together with the three arguments that they are *not* workaday person-likes surveyed earlier, and flexistentialism starts to look like a reasonable view.

11. The Argument from Coherent Wishes

Let me now develop the fourth argument for flexistentialism. Unlike the previous three, it does not appeal to (3) or the Counterfactual Comparative view of "bad for". The argument is less straightforward than the rest, but I find it the most interesting so I will develop it in a little more detail.

Those who reject the Counterfactual Comparative view typically agree that conceiving immediately was bad for Xia, but they explain this non-comparatively. For example, Shiffrin (1999) argues that it was bad for Xia insofar as it produced in her a condition that impedes the exercise of her agency, and Velleman (2008) says that it was bad for Xia because it violated her right to be born into good enough circumstances. This allows them to maintain that conceiving immediately was bad for Xia, regardless of whether she is better or worse off than she would have been had her parents waited. Grant that some account along these lines is right about the English term "bad for", so that the three previous

arguments for flexistentialism can be rejected. Still, I believe that all anti-flexistentialist views miss something out about nonidentity cases that flexistentialism captures easily.

To see this, consider the situation from Xia's standpoint. How might Xia feel about her parents' decision to conceive immediately? I am not asking how she might *morally* evaluate their decision; I am asking how she might feel from a *self-interested* perspective about what they did. I say it would be entirely reasonable of her to feel *angry*; in particular, angry *that they did not wait*. She might think: "I wish my parents had waited, so that I wouldn't be suffering so! I wish they had not been so hasty!"

But how can anti-flexistentialists make sense of this attitude? They think that Xia (the entity that matters) would not exist had her parents waited, so on their view Xia's wish amounts to a wish for something that would result in her never existing. But Xia does not wish that she never existed; she just wishes she weren't suffering! Nor *should* she wish she never existed: by stipulation of the case, she enjoys a worthwhile life. So, even if Shiffrin and company are right that conceiving immediately was bad for Xia in some sense, the fact remains that, according to anti-flexistentialism, Xia's wish is incoherent: it is a wish for a situation that she does not and should not wish for! Flexistentialism does better. It says that Xia would have been born had they waited, and so would have been better off had they waited. Her self-interested wish is then perfectly understandable: she is wishing for something that would have made *her* better off. Thus, the argument from coherent wishes is that flexistentialism best explains why Xia's attitude is reasonable.[22]

The argument is perhaps more compelling in the case of environmental planning. Recall Wanda, our descendant 300 years hence, struggling to eek out an existence because we did not conserve the natural environment. Surely she may reasonably wish that we had conserved. "If only they had conserved, I wouldn't have to struggle so! Would that they were less selfish!" The response that she would not exist if we had conserved, so that she should be *very glad indeed* that we were selfish, strikes me as philosophical in the pejorative sense—insensitive even. But that is what the anti-flexistentialist must say.

That is the rough idea behind the argument. It rests on three ideas:

(i) These wishes are coherent (this is the data).

(ii) Anti-flexistentialists views have trouble accounting for the data.

(iii) Flexistentialism easily accounts for the data.

Claim (iii) is unproblematic, but (i) and (ii) deserve more discussion. Regarding (i), the data is that the following attitudes are coherent:

> Xia's self-interested wish that her parents had waited.
>
> Wanda's self-interested wish that we had conserved the natural environment.

By "coherent", I do not mean rationally mandatory. Xia's condition might become an important part of her self-conception. She might find herself bonded to the community of people with her condition. As a result, she might reasonably come to prefer her life as it

[22] To be clear, the point is *not* that Xia's wish is the basis of a moral complaint against her parents. As Harman (2009) shows, inferences from retrospective wishes and regrets to moral or normative conclusions are deeply fraught. Partly for that reason, I avoid any such inference here.

is, with the condition, than a life without it, even if in some objective sense she would be better off without it.[23] Relatedly, Wanda may reflect on the fact that, if we had conserved and she were to somehow still exist, she would live an immeasurably different life with a different upbringing, a different family, a different career, and so on. So she may decide, on balance, that she would prefer the kind life she has, hardships and all, to a very alien life, even if the latter would be better for her in some objective sense.

These are what Wallace (2013) calls "reasons of attachment" to prefer one's actual life over others that would have been better in some objective sense, and I agree that these can be reasonable and decisive.[24] This is not to say that they *must* prefer their actual life: it could be that a vastly different life is precisely what Wanda wants, given how difficult her actual life is! Still, I agree that *if* Wanda had these attachments, it would be irrational for her to wish that we had conserved, and we might criticize her wish on those grounds. *Mutatis mutandis* for Xia.

What seems to me wrong is to criticize their wishes on the grounds that they are making a *metaphysical* mistake. This might be appropriate in other cases, for example when my daughter says that she wishes she were a dragon: I might reasonably respond that there is no coherent possibility in which she is a dragon and advise her to yearn for other things. But when Wanda wishes that we had conserved, it would be awfully obtuse to respond "But then you wouldn't exist! You should be *glad* that they were selfish and that you are consequently eeking out a meagre existence." Only a philosopher would say such a thing, and I do not mean that kindly! No, Wanda's attitude is the perfectly understandable wish that the past had been kinder to her, and should be treated as such. This is what I mean by saying that her wish is "coherent".

This data may be hard to see. As we have seen, many theorists have thought that claim (4)—that Xia would not exist if her parents had waited—is an obvious truism. As a result, they may have learned to bite the bullet on this point and think that Xia's wish is in fact incoherent, and that Xia should be *glad* that her parents conceived immediately after all. Indeed, I suspect that this bullet has been "internalized" to the point of being mistaken for data itself.[25] But I believe this is to let false theory blind one of the phenomena. With flexistentialism on the table, we see that (4) is not an obvious truism at all, but a piece of optional *theory*. To see the data, I must ask you to put aside your theoretical belief that Xia would not exist had her parents waited. We must see things afresh, from a perspective of theoretical innocence. Our question is then whether, intuitively, we should criticize (say) Xia's wish on metaphysical grounds and say that she must be glad to live her life of struggle. On the face of it, the answer is "no". To say "yes" strikes me as perversely over-intellectualized.

---

23 Harman (2009) discusses this kind of preference in some detail.

24 There is a growing literature on these kinds of preferences; for a partial guide, see Setiya (2014) and references therein.

25 Parfit (1984), for example, assumes that it is incoherent for future generations to wish that we had conserved the environment. This is evident when he argues that they would happily waive her right to be born into good enough circumstances: his point is based on the idea that they should hardly wish that we had done otherwise. Similarly, Liberto (2014), building on a case of Harman's (2004), appears to assume that it is incoherent to wish that someone has different biological parents. This is evident when she argues that a woman might end up glad she was raped on the grounds that the resulting child she loves would not have existed otherwise. The fact that she might be glad makes sense on the assumption that it is incoherent for her to wish that she had had the same child in another way.

Perhaps the data is most easy to see in yourself. If flexistentialism is false then your own existence hangs on a knife edge: it depends on your parents conceiving when they did, and *their* parents conceiving when they did, and so on back into the ages. Your existence then becomes an unimaginably improbable event, such that if history had differed at all significantly before your conception, you would not exist. If you are better off existing than not (which I imagine is true of many people reading this) it follows that it is incoherent for you to wish that history had been different. Thinking self-interestedly, you should adopt a somewhat Leibnizian view that the world up until your conception was pretty much the best of all possible worlds! I claim that this is clearly wrong. Suppose that 10 years before you were born, the government implemented a certain policy that had disastrous effects on the state school system. Surely you might wish, self-interestedly, that the policy had never been enacted, so that you would have enjoyed a better state education. But if flexistentialism is false, this wish is incoherent: we must criticize it as based on a metaphysical mistake, akin to the wish that one was born a dragon. My claim is that this ignores the clearly reasonable way in which people can wish that the past had been kinder to them.

So much for (i). Turn now to (ii), the claim that anti-flexistentialists have trouble explaining the data. When I quickly motivated it above, the central idea was this:

> (*) It is irrational for Xia to self-interestedly wish that her parents had waited, if that would have led to her non-existence.

For the anti-flexistentialist says that Xia would not exist if her parents waited. Hence, by (*), it follows that it is irrational for her to wish that her parents had waited.

But why think that (*) is true? Earlier I offered one reason: namely, that the case is stipulated to be one in which Xia lives a worthwhile life. We can also offer a second argument for (*), which appeals to the premise that Xia is *better off* existing with her condition than not existing at all. It then follows that Xia's wish is a *self-interested* wish for an event that would lead her *worse off*; and it seems plausible to conclude that it is therefore irrational. Now, I stress that this conclusion does not *logically* follow. As the discussion of reasons of attachment suggests, it can sometimes be rational to prefer scenarios in which one is worse off in some objective sense, so one might insist that Xia's wish is rational *even though* it is a wish for a situation in which she is worse off. Still, reasons of attachment arise when the worse situation is the *actual* situation, in which one has become attached to various projects or loved ones that would be absent in the better situation. This is not the case with Xia. Thus, while the argument is not valid, it nonetheless seems reasonable in Xia's case: *if* she is better off existing with the condition than not existing at all, then it does seem irrational for her to self-interestedly wish for a situation in which she would exist, as (*) states.

Now, this second argument for (*) will be rejected by those theorists who, like McMahan (1981), reject the intelligibility of comparing Xia's actual welfare to her welfare in a world in which she does not exist. To them, I offer the first argument. But I also press them to say how they can explain the coherence of Xia's wish. After all, such theorists typically say that conceiving immediately was bad for Xia in a non-comparative sense—say, by causing her various bads. If they are right, then by the same token it was also *good* for Xia in a non-comparative sense: it caused her various goods including life, some happiness, and so on. Why, then, would Xia be angry and *wish that her parents*

had waited? Why focus on the bads at the expense of the goods? Thus, regardless of (*), it is hard to see how these theorists can explain the coherence of Xia's wish; claim (ii) stands.[26]

This, then, is the "coherent wishes" argument: flexistentialism best explains the data that the wishes described here are coherent. Along with the three arguments discussed in section 9, it suggests that flexistentialism is not *just* a consistent view in logical space; it might even be true.

Before we move on, it is worth mentioning that some anti-flexistentialists are in a better position to explain Xia's wish than others. All anti-flexistentialists by definition agree that Xia would not exist had her parents waited. But as we saw in section 2, some anti-flexistentialists think it is *impossible* for Xia to exist if they waited, while others think her existence is (merely) *astronomically unlikely* conditional on them waiting. This second kind of anti-flexistentialist thinks that, strictly speaking, there is a small class of worlds C in which the couple waits and still have Xia, even though the counterfactual "If they had waited, Xia wouldn't exist" is true. So, these anti-flexistentialists could in principle argue that Xia's wish *that her parents had waited* is in fact a much more specific wish *that one of the worlds in C were actual*. Alternatively, they could try taking her wish at face value as the wish *that her parents had waited*, but argue that this is rational, contra (*), by detaching rational wishes from counterfactuals. For example, they could argue that, at the time of the couple's choice, the probability that Xia will exist conditional on their waiting was same low value as it was conditional on their conceiving immediately, so that (at the time of choice) Xia's expected utility conditional on their conceiving immediately was less than it was conditional on their waiting.[27] Perhaps *this* would explain why it is rational for her to prefer that they had waited. But I leave it for another time to assess whether these strategies are promising. For now, it suffices to say that the flexistentialist has a particularly simple explanation to hand: it is rational for Xia to self-interestedly wish that her parents had waited because *she would have been better off if they had*.

12. *De Dicto* Badness

One might suspect that the virtues of flexistentialism just surveyed are also enjoyed by another approach to the nonidentity problem. I have in mind the *de dicto* approach recently championed by Caspar Hare (2007).

Like others, Hare thinks that (4) is obviously true: if the couple had waited, Xia would not have been born. That being so, he agrees with (3) that being conceived immediately was not bad for Xia in a *de re* sense: she is better off than *she* would have been had they waited. Nonetheless, he says that it was bad for Xia in a *de dicto* sense: she is worse off than the entity x that would have satisfied the description "their first child" had they waited. True, x would not have been Xia, but that does not matter for *de dicto* badness; what matters is that x and Xia both satisfy the description "their first child". So long as we read (2) as allowing that an act might be wrong in virtue of being bad for someone in the *de dicto* sense, the puzzle is resolved.

---

[26] I am indebted to Johann Frick for this point. Note that the argument here is independent of the question of how we should weigh the goods and bads caused by an act, to determine whether the act was morally wrong.

[27] Roberts (2007, 2009) takes this line in some nonidentity cases.

Whether an act is bad for someone *de dicto* depends on a description. Being conceived immediately was bad *de dicto* for Xia relative to the description "their first child", as just explained. But suppose Xia was conceived on November 5th. Then being conceived immediately was *not* bad *de dicto* for Xia relative to the description "their child conceived on Nov 5th". This is because Xia, who actually satisfies that description, is *not* worse off than the entity *x* that would have satisfied that description had they waited, since there would have been no such entity *x*. This relativity to a description means that flexistentialism and the *de dicto* approach can mimic each other to some extent. When the *de dicto* view says that an act is bad *de dicto* for *x* relative to the description "the D", the flexistentialist can say that the act is *de re* bad for an entity that is essentially the D.

Nonetheless, the "coherent wishes" argument from the last section arguably tells against the *de dicto* approach. As we saw there, it seems reasonable for Xia to wish that her parents had waited so that she would not have to live a life of suffering. But how can the *de dicto* approach make sense of this? True, if they had waited, the entity *x* that would have satisfied the description "their first child" would have been free of pain. But this entity *x* would not be *Xia*. So, when thinking *self-interestedly* about what counterfactual situations she would prefer, Xia should *not* be concerned with *x*'s welfare but with her own—this is what "*self*-interest" means. Thus, like any approach that accepts (4), the *de dicto* approach implies that Xia's wish is incoherent, contrary to the data.

Here I assume that "identity matters" in self-interested counterfactual reasoning. Parfit coined this term when discussing self-interested *temporal* reasoning. He said that identity is not what matters, and his point was that I should care self-interestedly about any future entity connected to me in the right way, regardless of whether it is identical to *me*. Similarly, a fan of the *de dicto* approach might argue that when thinking self-interestedly about counterfactual scenarios, identity is not what matters: I should care about any counterfactual entity satisfying some privileged description D, regardless of whether it is me. I will not try to adjudicate this issue here. It suffices to say that flexistentialism preserves the platitude that identity matters—that *self*-interested reasoning about counterfactual scenarios involves concern about *me*—while the *de dicto* approach does not. But I leave it for another time to decide how weighty the platitude is.

A related problem for the *de dicto* approach arises with regard to third-personal notions like compensation. Suppose you think that Xia deserves compensation for what was done to her. If it helps, modify the case so that her parents decided to conceive immediately due to the egregious negligence of their doctor. Perhaps her parents voluntarily had their gametes screened to prevent conceiving a child with Xia's condition, but their doctor watched soccer instead of performing the tests and told them they were fine. Whatever the details of the case, the question is how to account for the fact that Xia deserves compensation. Now, I do not have a full analysis of compensation to hand, but let us assume that she deserves compensation in part because the doctor's negligence was bad for her. The flexistentialist can account for this easily: it was bad for her in the *de re* sense that she is worse off than *she* would have been had the doctor done the tests.[28]

---

[28] Of course, in compensating Xia we also compensate many coincident entities that do *not* deserve compensation. But this is no more puzzling than the fact that in buying a statue of Goliath we also buy a number of coincident entities that are not worth the price. Thanks to Sheldon Smith for pointing out this phenomena.

But how can the *de dicto* approach account for this? The idea is presumably that she deserves compensation (in part) because the doctor's negligence was bad for her in the *de dicto* sense. But remember, whether an act is bad *de dicto* for someone depends on a description. Relative to the description "their first child", the doctor's negligence was bad *de dicto* for Xia; relative to the description "their child conceived on Nov 5th", it was not. So, should Xia be compensated? There is no fact of the matter: she should under one description, but not under another. Thus on this approach the description-relativity of *de dicto* "bad for" induces a description-relativity on deserving compensation. And this is not what we wanted: the data was that Xia deserves compensation *simpliciter*. Indeed it is unclear what it could *mean* to deserve compensation under one description but not another. The physical mechanics of compensation mean that a person is either compensated or not—one can hardly compensate a description![29]

A fan of the *de dicto* approach might now say that there is a privileged description D, such that an act is bad for someone *simpliciter* iff it is bad for her relative to D. But even if an extensionally adequate account along these lines can be given, it is unclear how it might be morally justified. After all, why should *Xia* be compensated just because *someone else* would have been better off had the doctor done the tests? Of course, one might say with Shiffrin and Harman that the doctor's negligence was bad for Xia because it caused her pain. But the *de dicto* approach is an attempt to make sense of these phenomena in terms of counterfactual comparisons. The problem is that by using *de dicto* comparisons, rather than *de re* ones, it is hard to see how the comparisons carry moral weight regarding what *Xia herself* deserves.

13. Containing the Explosion

So far, I have made a case for flexistentialism on the assumption that there is an explosion of coincident entities. If we reject that assumption, does the case for flexistentialism fall away?

Recall how we got the explosion. As we saw in section 2, the nonidentity problem presupposes essentialism—claim (4) makes no sense unless there is an intelligible distinction between essence and accident. We then get an explosion of coincident entities with two assumptions:

> Assumption A: If essentialism is true, then ontic essentialism is true.
> Assumption B: If ontic essentialism is true, then unlimited essentialism is true.

So the question is whether, if we reject one of these assumptions, we can then reject flexistentialism and defend the standard approach to the nonidentity problem on which (4) is true of the entity that matters, so that (1)–(4) are straightforwardly inconsistent.

Note right away that Assumptions A and B are deeply metaphysical theses, usually ignored in discussions of the nonidentity problem. So, even if we defend the standard

[29] To be clear, this is not to deny that someone might deserve compensation *because* some description is true of them. Xia deserves compensation because she was born under certain circumstances, and not because (say) she had an ice cream on her 5th birthday. But this is not to relativize compensation to a description; it is just to say that she deserves compensation *simpliciter*, and that this is so because she was born into certain circumstances.

approach by denying these assumptions, we will have learnt something interesting about the metaphysical foundations of that standard approach.

Still, it turns out that defending the standard approach is harder than it might seem. Consider Assumption B. It says that if ontic essentialism is true, then *any* set of properties of Xia constitute the essential profile of *some* entity coincident with her. That is a strong assumption indeed! But we did not really use anything so strong. All we used is the idea that *some* sets of Xia's properties—in particular, sets that we had not previously focused on, like the set containing just the property of *being their first child*—constituted the essential profile of some entity. So, to deny what we used of Assumption B, one must defend a view on which there is just the entity that has its origins essentially, and maybe one or two other workaday person-like entities, but nothing else. The challenge that faces any such view is to draw some privileged distinction between the few sets of properties that *do* constitute the essential profile of something, and the sets of properties that do not. I do not claim that this is impossible, but it is unclear how it might be met. At the very least, it would be interesting to learn that the standard approach to the nonidentity problem rests on a solution to this problem.[30]

But in any case the point is moot, for Assumption B is a red herring. The core flexistentialist claim is that the practical person is not a workaday person-like entity. This is consistent with the *limited essentialist* view that the practical person is the *only* entity there—the flexistentialist can happily reject Assumption B! To be sure, Assumption B is a dialectically helpful ladder to use when convincing you that flexistentialism is a logically coherent view, for once we have all the coincident entities in place it is clear that one hypothesis is that *this* one matters, another hypothesis is that *that* one matters, and so on. But the ladder can be thrown away: even if the entity that the flexistentialist says matters is the only entity there, it remains coherent to assert that that is the entity that matters! The same goes with my arguments for flexistentialism. They did not depend on all the coincident entities actually being there; they just depended on our recognizing the coherence of different hypotheses about the essential properties of the things that matter.[31]

Thus, if one wishes to reject flexistentialism, there is little to be gained by rejecting Assumption B. That leaves Assumption A. It says that if essentialism is true, then ontic essentialism is true. Suppose we reject this assumption and become descriptive essentialists. What then does the nonidentity problem look like?

Remember that on descriptive essentialism, whether an entity has a given property essentially or accidentally depends on the manner in which the entity is described. Thus, whether it is true, of Xia, that *she* would not exist if her parents had waited, as (4) states, becomes relative to a description. Which description is relevant to evaluating (4)? The descriptive essentialist will presumably say that it is fixed by the moral context involved in talking about nonidentity cases. (She could in principle say that it is fixed by the kind of name used to formulate (4), but that would have the unattractive consequence that the ethical matters at stake in the nonidentity case depend on what *name* we use for the

[30] In this regard, note that we did not need to presume that the coincident entities are metaphysically on a par: it may be that the one with its origins essentially is more "fundamental" than the rest (whatever that means), so long as the rest are there. It is therefore not enough to argue that some of the coincident entities are "derivative" or "gerrymandered".

[31] I am extremely grateful to Gail Fine for helping me to see this point.

couple's child!) The question is then whether the relevant description fixed by context makes (4) true or false.

If we say that it makes (4) true, we have the standard approach to the nonidentity problem back again, this time in the language of descriptive essentialism, and our task will be to reject one of (1)–(3)—see the literature on the nonidentity problem to see how that goes. Alternatively, if we say that the relevant description makes (4) false, then we have a view that is ethically equivalent to flexistentialism. For flexistentialism says that that there may be many entities coincident with Xia, but the one that matters is not one that has its origins essentially. Correspondingly, this descriptive essentialist view says that there are many descriptions that pick out Xia, and the ones that matter in moral contexts—e.g. in discussions of the nonidentity problem—do not pick out her origins or genetics or time of conception. This view differs from flexistentialism under the hood, in its metaphysical workings, but ethically speaking they yield the same results. In particular, both insist, contrary to the standard approach, that if we focus on the entities/descriptions that matter, we should reject (4) and say that Xia would still exist if her parents had waited.[32]

Of course, the descriptive essentialist could instead say that (4) does not have a determinate truth-value, either because context does not supply a description to evaluate (4) with, or because the descriptions supplied do not fix a truth-value. But the result is a view that is vulnerable to the coherent wishes argument. For on this view there is no fact of the matter whether Xia exists in the counterfactual scenario in which her parents wait, and so there is no fact of the matter whether it is coherent of her to wish that they had waited, *contra* the data. The view is also vulnerable to my objection to the *de dicto* approach regarding compensation. Like the *de dicto* approach, this view implies that there is no fact of the matter whether Xia deserves compensation *simpliciter*: she deserves it relative to some descriptions, but not others, and that (on this view) is all there is to say about the matter.

Thus, to defend the standard approach to the nonidentity problem, it is not enough to reject Assumption A. One must also argue that the relevant descriptions that matter in moral contexts make (4) true. And here the four arguments presented in sections 9 and 11 can be translated into the language of descriptive essentialism to cast doubt on this idea. After all, it is not enough to say that (4) is "intuitively" true; one must argue that it is true relative to the morally important descriptions. The arguments in sections 9 and 11 constitute reasons to think that it is not.

In sum, it is true that I made two strong and controversial assumptions when motivating flexistentialism. But both are dispensable. Assumption B played no real role at all. And while Assumption A did, the dialectic discussed in this paper can be translated into a framework that rejects the assumption. Thus, complaining about the assumptions is not a productive way to defend the status quo.

14. Conclusion

Nonidentity cases are puzzling situations in which a choice affects not only the future distribution of welfare, but also the identities of the individuals over which that welfare

[32] One can interpret Meacham (2012) as providing such a theory. Though as he says on p. 265, he intends his proposal to be interpretable either as a descriptive essentialist view, or as a version of the *de dicto* approach.

is distributed. These cases have been widely discussed both in the abstract and in application to questions of disability rights, environmental conservation, reparations, and other topics. Almost across the board, these discussions assume a *metaphysical* claim about the essential properties of the entities that matter, and on that basis draw a variety of *ethical* conclusions about what makes various actions *wrong*, or the conditions under which an act is *bad for* someone, and so forth. Here I developed an alternative view on which the metaphysical claim assumed in these discussions is false. On this alternative view, the above description of the cases—that the choice affects the identities of the individuals involved—is true of the workaday person-like entities, but not of the entities that matter. Establishing that this alternative view is coherent is enough to show that the ethical conclusions standardly drawn from these examples are less motivated than we thought. But I also presented four considerations in favor of this alternative view. We cannot, of course, conclude that this alternative view is correct—that would require a book-length comparison of the various views on offer. But I do hope to have shown that it deserves our consideration.[33]

References

Bennett, Karen. (2004). "Spatio-Temporal Coincidence and the Grounding Problem". *Philosophical Studies,* 118(3): 339–371.

Breton, André. (1987). *Mad Love*. Translated by Mary Ann Caws. Lincoln, NE: University of Nebraska Press.

Burke, Michael. (1994). "Preserving the Principle of One Object to a Place: A Novel Account of the Relations among Objects, Sorts, Sortals and Persistence Conditions". *Philosophy and Phenomenological Research*, 54(3): 591–624.

Bykvist, Krister. (2007). "The Benefits of Coming into Existence". *Philosophical Studies*, 135(3): 95–115.

Fine, Kit. (1994). "Essence and Modality". *Philosophical Perspectives*, 8: 1–16.

——. (1999). "Things and Their Parts". *Midwest Studies in Philosophy*, 23(1): 61–74.

Freeman, Michael. (1997). *The Moral Status of Children: Essays on the Rights of Children*. The Hague: Kluwer.

Hanser, Matthew. (1990). "Harming Future People". *Philosophy and Public Affairs*, 19 (1): 47–70.

Hare, Caspar. (2007). "Voices from Another World: Must We Respect the Interests of People who do not, and Will Never, Exist?" *Ethics*, 117(3): 498–523.

Hare, Caspar. (2013). *The Limits of Kindness*. Oxford: Oxford University Press.

Harman, Elizabeth. (2004). "Can we Harm and Benefit in Creating?" *Philosophical Perspectives*, 18: 89–113.

——. (2009). ""I'll Be Glad I Did It" Reasoning and the Significance of Future Desires". *Philosophical Perspectives*, 23: 177–199.

Harris, John. (1998). "Rights and Reproductive Choice". In John Harris and Soren Holm (eds.), *The Future of Human Reproduction: Ethics, Choice and Regulation*. Oxford: Clarendon Press.

[33] Thanks to Elizabeth Barnes, Johann Frick, Laura Gillespie, Caspar Hare, Liz Harman, Louis-Philippe Hodgson, David Kovacs, Hallie Liberto, John Morrison, David Plunkett, and Melinda Roberts for their insightful feedback and encouragement while composing this essay. Special thanks to Laura Gillespie for suggesting the epigram.

Heyd, David. (1998). "Procreating and Value: Can Ethics Deal with Futurity Problems?" *Philosophia*, 18(2): 151–170.
Johnston, Mark. (2006). "Hylomorphism". *The Journal of Philosophy*, 103(12): 652–698.
Koslicki, Kathrin. (2008). *The Structure of Objects*. Oxford: Oxford University Press.
Lewis, David. (1986). *On the Plurality of Worlds*. Oxford: Blackwell.
McMahan, Jeff. (1981). "Problems of Population Theory". *Ethics*, 92(1): 96–127.
Meacham, Christopher J. G. (2012). "Person-affecting Views and Saturated Counterpart Relations". *Philosophical Studies*, 158(2): 257–287.
Parfit, Derek. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
Roberts, Melinda. (1998). *Child vs Childmaker: Future Persons and Present Duties in Ethics and the Law*. Lanham, MD: Rowman and Littlefield.
——. (2007). "The Nonidentity Fallacy: Harm, Probability and Another Look at Parfit's Depletion Example". *Utilitas*, 19(3): 267–311.
——. (2009). "The Nonidentity Problem and the Two Envelope Problem: When Is One Act Better for a Person than Another?". In Melinda Roberts and David T. Wasserman (eds.), *Harming Future Persons: Ethics, Genetics and the Nonidentity Problem*. Springer Dordrecht Heidelberg: London.
Savulescu, Julian. (2002). "Deaf Lesbians, "Designer Disability", and the Future of Medicine". *BMJ*, 325: 771–773.
Setiya, Kieran. (2014). "The Ethics of Existence". *Philosophical Perspectives*, 28: 291–301.
Sher, George. (1981). "Ancient Wrongs and Modern Rights". *Philosophy and Public Affairs*, 10(1): 3–17.
——. (2005). "Transgenerational Compensation". *Philosophy and Public Affairs*, 33(2): 181–200.
Shiffrin, Seana Valentine. (1999). "Wrongful Life, Procreative Responsibility, and the Significance of Harm". *Legal Theory*: 117–148.
——. (2009). "Reparations for U.S. Slavery and Justice Over Time". In Melinda A. Roberts and David T. Wasserman (eds.), *Harming Future Persons: Ethics, Genetics and the Nonidentity Problem*. Springer Dordrecht Heidelberg: London.
Thompson, Janna. (2001). "Historical Injustice and Reparation: Justifying Claims of Descendants". *Ethics*, 112(1): 114–135.
Velleman, David J. (2008). "Persons in Prospect". *Philosophy and Public Affairs*, 36(3): 221–288.
Wallace, R. Jay. (2013). *The View from Here: On Affirmation, Attachment, and the Limits of Regret*. Oxford: Oxford University Press.
Wasserman, David. (2005). "The Nonidentity Problem, Disability, and the Role Morality of Prospective Parents". *Ethics*, 116(1): 132–152.
Wolf, Clark. (2009). "Do Future Persons Presently Have Alternate Possible Identities?". In Melinda A. Roberts and David T. Wasserman (eds.), *Harming Future Persons: Ethics, Genetics and the Nonidentity Problem*. Springer Dordrecht Heidelberg: London.
Woodward, James. (1986). "The Nonidentity Problem". *Ethics*, 96(4): 804–831.
Yablo, Stephen. (1987). "Identity, Essence, and Indiscernibility". *Journal of Philosophy*, 84(6): 293–314.