Critical Notice: <u>Child versus Childmaker: Future Persons and Present Duties in Ethics and the</u>

<u>Law</u>, Melinda Roberts, 1998, Rowman & Littlefield. *Nous* 34 (2000): 634-47. (Formerly entitled

"On the moral status of future people".)

PETER VALLENTYNE

1. Introduction

In <u>Child versus Childmaker</u> Melinda Roberts provides an enlightening analysis and a cogent

defense of a version of the person-affecting restriction in ethics. The rough idea of this restriction

is that an action, state of affairs, or world, cannot be wrong, or bad, unless it would wrong, or be

bad for, someone. I shall focus solely on Roberts's core principles, and thus shall not address her

interesting chapter-length discussions of wrongful life cases and of human cloning cases.

The person-affecting intuition can be spelled out in a deontic and in an axiological form:

<u>Deontic Person-Affecting Restriction</u>: An action, state of affairs, or world is <u>wrong</u> only if it

would wrong someone.

<u>Axiological Person-Affecting Restriction</u>: An action, state of affairs, or world is <u>worse</u> than

another only if it is worse for someone.

Roberts develops and defends the deontic version of the person-affecting restriction.

Indeed, as her discussion (and those of others) makes clear, it is quite unlikely that there is a

coherent version of the axiological form of the intuition (e.g., because of problems of

transitivity). Roberts rightly rejects any appeal to <u>impersonal</u> ranking of worlds, and appeals only

to the (many) personal rankings of worlds of the individuals involved.

Roberts defends the view that worlds in which a person <u>does not exist</u> can be ranked in terms of that person's well-being (how good that world would be for him/her) along with the worlds in which he/she does exist. Worlds in which the person has a life worth living are ranked more highly for that person than worlds in which he/she doesn't exist (along with some worlds with indifferent existence), and the latter worlds are ranked more highly than worlds in which the person exists but doesn't have a life worth living. Although this is somewhat controversial (some would deny that we can assess how good a world is for a person who doesn't exist in it), it seems exactly right to me. Throughout I shall assume, as does Roberts, that non-existence has a value of zero (and thus that lives worth living have positive values, and that lives not worth living have negative values).

In what follows I shall formulate Roberts's theory as a theory of the permissibility of <u>actions</u>. In most of the book, Roberts formulates her theory as one of the permissibility of <u>worlds</u>, but, as I shall argue below, such assessments have little normative relevance, since they ignore the probabilities of realization (which depend on what actions are performed). The core of Roberts's theory is just as plausible, and in some ways more powerful, when (as is sometimes the case in the book) it takes actions to be the objects of assessment.

To start with, however, I shall make a highly restrictive assumption (not made by Roberts) that will effectively eliminate the difference between assessing actions and assessing worlds within the sort of welfaristic consequentialist framework that Roberts presupposes. I shall assume that if a given action were performed, then a specific world would be realized.[1] We shall assume, that is, that there are no uncertainties involved. Everything is completely determined by which of the feasible actions is performed. Later in this paper, I will address the problems that arise once this simplifying assumption is dropped.

2. Personalism Introduced

The Deontic Person-Affecting Restriction asserts that an action is wrong if it would wrong someone, but not otherwise. Roberts defends the following version of this principle:

P* (Permissibility): Action X is permissible at (time) t for (agent) s iff no one would be wronged by X, except perhaps people who in this choice situation died prior to t.[2]

P* rules out impersonal wrongs. In addition, it presupposes either that dead people cannot be wronged, or that if they can be wronged, such wronging is irrelevant to permissibility. Some who accept the general Deontic Person-Affecting Restriction may therefore reject P* on the grounds that dead people can be wronged and such wronging makes an action wrong. In what follows, however, I shall grant this aspect of P*.

On its own P* rules out little of substance. For almost any theory of permissibility can satisfy P* if an appropriate theory of personal wronging is cooked up. Consider, for example:

UPW (Utilitarian Personal Wronging): A person would be wronged by action X (at t by s) just in case (1) X would not maximize the total [or: average] well-being, and (2) X makes him/her is less well off than an alternative that would maximize the total [or: average] well-being.

Total utilitarianism satisfies both P* and the "total" version of UPW. No one would be wronged by actions that would maximize total well-being, and at least one person would be wronged by actions that would not maximize this total. Likewise, average utilitarianism satisfies P* and the "average" version of UPW.

Without some criteria of when a person would be wronged, P* has little bite. P* provides the framework, but the substance comes from the principles of personal wronging. Roberts supplies a partial set of criteria: two necessary conditions for wronging and one sufficient condition. One is:

M* (Maximal Well-Being): s would not be wronged by the performance of X if the performance of X would produce at least as much well-being for s as the performance of any feasible alternative would.[3]

Those who believe that there can be harmless wronging—such as harmlessly violating someone's property rights, or rights-violating paternalism—will reject this condition, but within the context of optimizing welfarist consequentialism it is highly plausible.

Indeed, M* is not a condition particular to person-affecting approaches. Both total and average utilitarianism, formulated in terms of P* and UPW, satisfy M*. If an individual has her maximal feasible well-being, then she does not have less well-being than under some alternative that maximizes the total, or average, well-being. Hence she is not wronged according to UPW.

A second principle of personal-wronging defended by Roberts is:

N* (Non-Existence): s would not be wronged by the performance of X if s would never exist if X were performed.[4]

Here and throughout we follow Roberts and appeal to existence in a way that is not time-indexed. Thus, the people who exist only before, or only after, an action is performed exist in this sense if the action is performed—as long as they definitely exist at some point if the action is performed. So N* does not require that the wronged individual exist at the time of action.

Consider, for example (where there are just two people, just two feasible actions, and "*" denotes non-existence):

	Smith	Jones
X:	2	*
Y:	1	5

Example 1

Here M* tells us that Smith would not be wronged by X and N* tells us that Jones would not be wronged by X. Hence P* tell us that X is permissible. M* also tells us that Jones would not be wronged by Y, but the principles so far are silent about whether Smith would be wronged by Y.

Unlike M*, N* is rejected by both the total and average versions of utilitarianism in the form of P* and UPW. In the previous example, both total and average utilitarianism judge X impermissible and Y obligatory (when they are the only two options). In their UPW form they each hold that Smith would not be wronged by X and that Jones would not be wronged by Y

(since, in each case, the person is maximally well off). In order to judge X impermissible they need therefore to hold that Jones (the only other person) would be wronged by X even though she wouldn't exist if X were performed, and that violates N*. Only by allowing people who would not exist to be wronged can total and average utilitarianism be squeezed into P* and UPW. So, unlike M*, N* is distinctively person-affecting in nature.

I take it that N* is plausible. People who would not exist if an action were performed would not be wronged by that action.

Roberts's final principle is:

D* (Deprived Gratuitously): s would be wronged by the performance of X if s would exist if X were performed and there is a feasible alternative Y such that:

(i) s would have more well-being if Y were performed than if X were performed;

(ii) no person who would exist if X were performed would also exist, but with less well-being, if Y were performed; and

(iii) everyone who would exist if Y were performed would also exist if X were performed.[5]

The rough idea of D* is that a person would be wronged if she would be "gratuitously" made worse off than she would be under some alternative. More exactly, a person would be wronged by X if some alternative would make her better off, that alternative would not make any of the other people who exist under X have a worse existence (i.e., both exist and have less well-being), and that alternative would introduce no new people.

In the above example, D* judges that Y would wrong Smith. For Smith could be made better off (under X) without creating any additional people and without making anyone else have a worse existence. It's important to note that Jones would be <u>worse off</u> under X than under Y

(since non-existence is worse for him than a life worth living of five), but she would not have a worse existence (since she would not exist if X were performed).

D* is a kind of optimizing condition. Those who reject either welfarist consequentialist theories of permissibility (e.g., various rights-based theories), or who reject optimizing versions thereof (e.g., satisficing, strict retributivist, or strict egalitarian theories), will likely reject this condition. Within the context of optimizing welfarist consequentialist theories there is something right about this principle. Nonetheless, D* is, I shall argue below, the principle that is most vulnerable to rejection within a person-affecting approach.

Personalism is the name of Roberts's preferred version of the person-affecting intuition, and is defined as the conjunction of P*, N*, M*, and D*. It is not a complete theory, since it does not provide necessary and sufficient non-moral conditions for permissibility. Some complementary theory of when a person would be wronged by an action (that appeals, for example, to rights or to equality promotion) is needed fill in the gaps left open by N*, M*, and D*. Roberts leaves the nature of the complementary theory open (although she tentatively favors an appeal to equality).[6]

Personalism avoids problems that totalism and averagism confront. First, it denies (via N*) that it is ever wrong not to procreate simply because it deprives a possible being of life worth living (since only beings that would exist can be wronged by an action). If it is wrong, it must be because some of the beings that would exist are wronged (e.g., gratuitously deprived of well-being). Consider, for example:

	Smith	Jones
X:	10	*
Y:	10	20

Example 2

If X and Y are the only two feasible alternatives, then both totalism and averagism judge X wrong (since it has a lower total and a lower average than Y). Personalism, on the other hand, holds that Smith would not be wronged by X (by M*, given that 10 is Smith's maximum) and that Jones would not be wronged by X (by N*, given that Jones would not exist if X were performed). By P* X is permissible (since no one would be wronged by X).

Personalism also avoids the Repugnant Conclusion that totalism faces. In its standard version the Repugnant Conclusion holds that a world in which everyone has a flourishing life is morally (impersonally) <u>worse</u> than a world in which there are sufficiently many more people each with a life just barely worth living (but with greater total well-being). Personalism avoids this version of the Repugnant Conclusion by not invoking any notion of impersonal or moral goodness. It appeals only to how good worlds are for each individual (a three place relation rather than a two place relation). Hence, it does not involve the repugnant conclusion in this form.

A different version of the Repugnant Conclusion concerns permissibility rather than moral goodness. It states that it is impermissible to perform an action that produces a world in which everyone has a flourishing life when there is a feasible alternative that produces a world in which there are sufficiently many more people each with a life just barely worth living (so that the total is greater). Totalism has this implication. Personalism, however, does not. Furthermore, Personalism is appropriately sensitive to whether the people who would exist if one action were performed are entirely different (with no overlap) from the people who would exist if the other action were performed. For example, suppose that there are just two feasible actions as follows:

Smith Jones1 Jones2 Jones3 …   Jones7 Jones8 Jones9

| X: | 8 | * | * | * | … | * | * | * |
| Y: | * | 1 | 1 | 1 | … | 1 | 1 | 1 |

Example 3

Here totalism says that X is impermissible and that Y is obligatory (since Y has a greater total). Personalism, on the other hand, says that each is permissible. This is because each action is such that the people it would produce are as well off as they can be (and thus no one would be wronged according to M*, and the action is permissible according to P*). The implication that Y is optional is not entirely unproblematic (we'll return to this point below), but it is significantly less problematic than totalism's judgement that Y is obligatory. Furthermore, Personalism takes a different view of the matter if some of the people affected are the same under both actions. Consider:

	Smith	Jones1	Jones2	Jones3	…	Jones7	Jones8
X:	8	*	*	*	…	*	*
Y:	1	1	1	1	…	1	1

Example 4

Totalism treats this case just as it treats the previous case, and judges Y obligatory. Personalism, on the other hand, treats this case differently. It says that X is obligatory and Y is impermissible. Y is impermissible because Smith is, according to D*, gratuitously less well off under Y than under X. X makes Smith better off, none of the people who exist under Y have a worse existence

9

under X, and everyone who exists under X (viz. Smith) exists under Y. Hence M* holds that Y would wrong Smith, and P* accordingly judges Y impermissible.

Let us now examine some of the key weaknesses of Personalism.

3. Personalism Criticized

M* and N* are, I have suggested quite plausible in the context of an optimizing welfarist consequentialist framework. D*, I shall now argue, may be a bit too strong. I shall also raise some challenges to Personalism taken as a whole, but I shall suggest that these challenges can be met.

3.1 Assessing D*

D* may be too strong. Consider:

	Smith	Jones	Frost
X:	3	3	3
Y:	5	*	3
Z:	*	5	3

Example 5

Here M* and N* rightly say that no one is wronged in Y or Z, and P* rightly concludes that Y and Z are permissible. Consider, however, X. D* says that Smith would be wronged by X (since she can be made better off under Y without anyone having a worse existence), and P* then concludes that X is not a permissible action. D* always judges it wrong to add a new person to

10

the world (e.g., Jones under X), when instead one could increase the well-being of someone (e.g., Smith under Y) who would exist in any case if the new person were added. The fact that X has a greater total, and is more equal, than Y is deemed irrelevant by D*.

A weaker person-affecting approach could reject the view that it is automatically wrong to bring a person into existence instead of benefiting someone who would exist in any case if that person were brought into existence. More specifically, D* could be weakened as follows so as not to require that X be judged impermissible here.

D**: s would be wronged by X if s would exist if X were performed and there is a feasible alternative Y such that:

(i) s would have more well-being if Y were performed than if X were performed;

(ii) no person who would exist if X were performed would be <u>worse off</u> if Y were performed; and

(iii) everyone who would exist if Y were performed would also exist if X were performed.

D** is simply D* but with clause (ii) strengthened to require that no one be <u>worse off</u>, as opposed to that no one have a <u>worse existence</u> (as in D*). Unlike D*, D** holds that it may sometimes be permissible to perform an action that would bring new people into existence even though one could alternatively bring only some of these new people into existence and make some of them better off. In the above case, for example, M* ensures that neither Jones nor Frost would be wronged by X, but N* and D** (unlike D*) are each silent about whether Smith would be wronged by X. Thus, the complementary theory must answer the question of whether they would be wronged here. And the answer it gives may depend, for example, on whether it is concerned with the average, total, or equality, in such cases.

11

Unlike D*, D** does not hold that making <u>people happy</u> takes absolute priority over making <u>happy people</u>. Most theorists attracted to person-affecting approaches will, I suspect, endorse this priority. My point here is simply that a weaker version of Personalism is possible. Furthermore, D** joined with M*, N*, and P* does not reintroduce the Repugnant Conclusion in its strong form. For like the D* version of the theory, the D** version holds that it is permissible to choose worlds with fewer people each of who is maximally well off. The only issue is whether it is also sometimes permissible to choose worlds with more people, when, for example, these worlds have a suitably greater total or more equal distribution. Some theorists may wish to pursue this weaker person-affecting alternative.

3.2 Assessing Personalism as a Whole

The main problem confronting Personalism as a whole is the infamous <u>non-identity problem</u> (emphasized by Parfit and others) that all person-affecting approaches face. According to Personalism, a person cannot be wronged if there is some feasible action such that (1) she will not exist if the action is not performed and (2) she will exist with a life worth living if that action is performed. For if any other action is performed, then the person will not exist and thus by N* is not wronged. And if the particular action is performed, then the person will have positive and maximal feasible well-being, and thus by M* is not wronged. (I continue to assume here that there is no uncertainty concerning the outcomes of actions. This assumption will be dropped in the next section, and things then become a bit more complex.) Given the precariousness of existence, this has broad implications. Even slight changes of action can lead a woman's egg to combine with a different sperm (of which there are typically 100-300 million in a typical male ejaculate!) or not to combine at all. Thus, it seems, most of the people who would come into

existence if a given action were performed would not have existed if a different action had been performed. Consequently, whatever positive well-being such a future person has (relative to the act in question), this well-being is, it seems, as great as possible for that individual. Thus, M*, it seems, holds that the individual would not be wronged.

Roberts defends the view (entailed by M*) that, in cases where a person exists only if a specific action is performed, that person is not wronged by that action if she has a life worth living. She rightly points out, however, that such cases are highly artificial. In most cases that we confront or discuss, for most feasible actions there are many people who would exist if the action were performed who would also exist if some other action were performed. Such people can be wronged.

Consider, for example, Gregory Kavka's pleasure pill example. Here a man and a woman each take a pill that slightly increases their sexual pleasure and that induces mild retardation in any child conceived. Conception, we'll suppose, is a certainty if they have sex. Furthermore, pausing to take the pill has the result that a different sperm combines with the egg, and thus changes who exists. The situation thus might look like this:

	Mother	Father	Child1	Child2
No Sex:	9	9	*	*
No Pill Sex:	10	10	10	*
Pill Sex:	11	11	*	5

Example 6

If these are the only three feasible alternatives, then Personalism says that Pill Sex (producing a mildly retarded child) is permissible. For Mother, Father, and (retarded) Child2 are each

maximally well off under this action, and thus by M* they are not wronged. Child1 does not exist, and thus by N* is not wronged. Thus no one would be wronged and P* judges Pill Sex permissible.

Roberts rightly holds that if there are only the three portrayed alternatives, then Pill Sex is permissible in this situation. She is highly skeptical, however, that any real life situation, or any of the usual hypothetical cases discussed, sex is like this. For there is always going to be at least one additional feasible action: The parents could pause as if they were taking the pill—thereby creating the conditions for the conception of Child2—and then have sex. The choice situation might thus look like:

	Mother	Father	Child1	Child2
No Sex:	9	9	*	*
No Pause No Pill Sex:	10	10	10	*
Pause No Pill Sex:	10	10	*	10
Pill Sex:	11	11	*	5

Example 7

In this situation, Personalism leaves open whether Pill Sex is permissible. For it does not tell us whether Child2 would be wronged by Pill Sex. The complementary theory will need to address this issue, and it may well judge that Child 2 would be wronged by Pill Sex (e.g., by appealing to the greater equality of Pause No Pill Sex).

So in most real life cases, for any given person there is typically going to be more than one action which brings him/her into existence, and the action that is better for that person may be worse for someone else who exists under both actions. In such cases, Personalism will appeal

14

to its complementary theory to assess who would be wronged.

So, in fact, Roberts has insightfully pointed out, existence is not so precarious that every contingently possible future person will exist only if some particular action is performed. Still, it is at least possible in principle that all or most such people will exist only if some particular action is performed, and we need to assess Personalism in such cases. More specifically, we need to assess Personalism's claim that an action that produces only people who would only exist if that action were performed is permissible as long as each person has a life worth living.

To help test this view, consider the following two cases:

	Smith	Jones1	Jones2	Jones3	…	Jones1M
X:	1	*	*	*	…	*
Y:	*	1M	1M	1M	…	1M

Example 8

	Smith	Jones	Black	White
W:	1	17	*	*
Z:	*	*	9	9

Example 9

In each of these two cases, Personalism judges both of the feasible actions permissible, since for each action every existing person is as well off as he/she can be. And yet, in Example 8 the second action (Y) is vastly superior in terms of aggregative efficiency (total well-being), and in Example 9, the second action (Z) is vastly superior in terms of equality. Is the first action of each

case really permissible as claimed by Personalism?

The issue is not whether aggregative efficiency or equality is a relevant consideration for determining permissibility. Personalism is compatible with a complementary theory that makes them relevant. The complementary theory, however, applies, only where M*, N*, and D* are silent. In the above two examples, those principles leave no room for the complementary theory, since, in the context of P*, they fully determine what is permissible. The issue here is whether aggregative efficiency, equality, or some other consideration ever generates an <u>impersonal</u> wrong. I'm inclined to think that Roberts is right to endorse P* and thus reject impersonal wrongs, but many people will be unconvinced. The issue is, of course, highly complex, and space limitations prevent me from pursuing this issue further.

4. Contingent Existence Relative to an Action

We have been assuming that if a given action is performed a specific world will be realized. We have been assuming, that is, that there is no uncertainty about what will happen if a given action is performed. This highly unrealistic assumption (made by me, not by Roberts) was made so as to sidestep the difference between assessing the permissibility of worlds as Roberts does, and assessing the permissibility of actions. It is now time to reexamine Personalism once this assumption is dropped.

To see the problem that arises once the simplifying assumption is dropped, suppose that I can perform one of two actions. If I perform X, then with probability .99 Jones will have an absolutely horrible life of -100, and with probability .01 she will have a wonderful life of 100. My only alternative is to perform Y, in which case Jones has with certainty an almost wonderful life of 99. No one else (including myself) is affected by what I do. I perform the first action, but

16

fortunately—despite its very low probability—Jones has a wonderful life. Have I wronged Jones?

On a natural interpretation of M*, I have not. For Jones has the best possible life that she could have, and M* could be interpreted as saying that such people are not wronged. But surely I have wronged her. I have exposed her to a high probability of a horrible existence for only a small chance of a small gain. Of course, as things turned out, things are quite nice. But that does not detract from the fact that I did something wrong. If I fire a loaded gun at her head and by good luck miss, I have still wronged her.

How can Personalism, and person-affecting approaches in general, deal with this sort of case? One approach—that favored by Roberts—is to formulate Personalism as a theory of permissibility for worlds and not for actions. (Recall that her official formulations of her principles are in terms of the permissibility of worlds, and that I reformulated them in terms of the permissibility of actions.) The permissibility of a world relative to a given time is determined by how it compares with the other worlds that are empirically possible at that time. Personalism so construed says that the world in which I perform an action that makes it highly likely that Jones will have a horrible existence, but in which Jones has (unlikely as it was) a wonderful life, is a permissible world relative to the alternative empirically accessible worlds at that time. After all, no accessible world is better for Jones. The problem with this approach is that it totally ignores the probabilities of things turning out otherwise. It just looks at the end result. This would be fine if one could choose at will to bring about these end results (thereby eliminating the possibility of things turning out otherwise), but given that one can't do so, assessing the permissibility of worlds in this way is quite irrelevant. The assessment of actions is where the real action is.

A closely related approach is to assess the permissibility of actions but to hold that the permissibility of actions depends solely on how things happen to turn out (and does not depend

on the possibilities and probabilities of things turning out otherwise). On this approach I do nothing wrong in exposing Jones to the high probability of a horrible existence, as long as it turns out that Jones has a wonderful life. Again, the problem is that it ignores the very real and relevant possibilities and probabilities. It was lucky for Jones that he had a wonderful life, but such luck is irrelevant to the assessment of my action, since it is the result of factors beyond my control. I control whether Jones is exposed to the risk of a horrible existence, and I inappropriately exposed him to this risk. My action is thus wrong, even though due to factors beyond my control things turned out rather nicely.[7]

The only plausible way for Personalism to deal adequately with (the very common) cases where there is uncertainty about what will happen if an action is performed is to appeal to some measure of well-being for people that is sensitive to the various possibilities and probabilities. One natural way to do this is by taking the expected value, and for simplicity I will assume well-being is understood as expected well-being. So interpreted, M* rightly says that my action that exposed Jones to the high probability of a horrible existence wrongs her.

There is one other problem that we have to consider now that we recognize that there is uncertainty in what will happen if a given action is performed. It concerns how we treat non-existence. To see the problem, suppose that I can perform one of two actions. If I perform X, then with probability .99 Jones will come into existence with an absolutely horrible life of -100, and with probability .01 Jones will not exist (and thus have 0 well-being). My only alternative is to perform Y, in which case Jones will definitely exist and will definitely have a very enjoyable life of 10. No one else (including myself) is affected by what I do. Unlike the previous example, in this example Jones may, but also might not, exist if I perform the first action. I perform the first action, but fortunately—despite the very high probability of a horrible existence—Jones does not come into existence. Have I wronged Jones?

On a natural interpretation of N*, I have not. For Jones does not exist, and N* seems to say that people who do not exist are not wronged. But surely I have done something wrong here. I have exposed a possible person to a high probability of a horrible existence for no good reason. Of course, as things turned out, things aren't so bad. It's much better that Jones not exist than that he have the horrible existence. But that does not detract from the fact that I did something wrong.[8]

The only plausible way for Personalism to deal adequately with (the very common) cases where there are people who may, but may not, come into existence if a given action is performed, is to treat them on a par with people who will <u>definitely</u> exist if the action is performed, but with their well-being reflecting the possibility and probability of non-existence. N* on this approach does not apply. It applies only to people who will <u>definitely</u> <u>not exist</u> if the action is performed.

This approach requires that D* should be revised to (changes are emphasized):

D*** (Deprived Gratuitously): s would be wronged by the performance of X if s would exist if X were performed and there is a feasible alternative Y such that:

(i) s would have more (expected) well-being if Y were performed than if X were performed;

(ii) no person who <u>might</u> exist if X were performed (a) <u>might</u> also exist if Y were performed, <u>and</u>

     (b) <u>has more (expected) well-being under Y than under X</u>; and

(iii) everyone who <u>might</u> exist if Y were performed <u>might</u> also exist if X were performed.

This revision is needed to capture the idea that those who might exist if a given action were performed have the same status as those who would exist.

This approach is, I believe, the only plausible way for Personalism to go. We should note, however, that treating in this way people who are merely possible if a given action is performed,

reduces the distinctive bite of Personalism. For, as we have seen, the distinctive nature of Personalism comes primarily from its treatment of non-existence in N* and in D*. By reserving this treatment of non-existence for those who will <u>definitely</u> not exist if a given action is performed, the distinctive treatment is more limited.

5. Conclusion

By and large Roberts's Personalism seems exactly right. I suggested, however, that D* may need to be weakened to D** to leave more room for total or equality considerations. And I argued that the theory needs to be formulated as theory of permissibility of actions (as opposed to worlds). Once so formulated, the only plausible way to go is to treat those who might, but need not, exist if an action is performed the same way as those who will definitely exist. These are relatively small points compared to the very significant advance that Roberts has made in the formulation of plausible person-affecting principles. All future discussion of the person-affection intuition will need to take her work into account.[9]

Notes

[1] Throughout, subjunctive conditionals are to be understood as follows. Where A is empirically possible relative to the time of assessment, "If A were to obtain, then B would obtain" holds just in case in <u>all</u> empirically possible worlds (of which there may be many), every A-world is a B-world.

[2] All formulations of Roberts's principles are mine, and are in terms of the permissibility of actions rather than of worlds. Roberts's official formulations are in terms of worlds and will be given in footnotes. Her official formulation of P* is: X is permissible for agents at t if and only if no person <u>who exists at or after t</u> is wronged in X at or after t. Because the emphasized phrase of this principle introduces a redundancy given her principle N* (which says that non-existents cannot be wronged), I have for expository reasons weakened P* to read "X is permissible for agents at t if and only if no person is wronged in X at or after t, <u>except perhaps people who in X died prior to t</u>. In the context of her principle N*, the two versions are equivalent.

[3] Roberts's official formulation of M*: s is not wronged by agents in X if, for each world Y accessible to such agents, s has at least as much well-being in X as she has in Y.

[4] Roberts's official formulation of N*: s is not wronged by agents in X if s never exists in X.

[5] Roberts's official formulation of D*: s is wronged by agents in X if s exists in X and there is some world Y accessible to such agents such that: (i) s has more well-being in Y than s has in X; (ii) for each person s' who ever exists in X either: s' has at least as much well-being in Y as s' has in X, or s' never exists in Y; and (iii) for each s' who ever exists in Y, s' exists at some time in X. In n. 48 of Ch. 2 Roberts identifies a possible strengthening of this principle, which she has endorsed in personal correspondence. The strengthening involves replacing clause (iii) with the

following (new parts emphasized): (iii') for each s' who ever exists in Y, <u>either</u> s' exists at some time in X <u>or s' is maximally well off in Y.</u> Although I agree that this is a plausible strengthening, I shall ignore it in what follows.

[6] It's worth noting that Roberts's principles do not require that well-being be cardinally measurable or interpersonally comparable. The principles appeal only to intrapersonal ordinal well-being. Of course, a plausible complementary theory will need to appeal to more than this in order to deal with tradeoffs among people.

[7] I develop this criticism in Peter Vallentyne, "Utilitarianism and the Outcomes of Actions", <u>The Pacific Philosophical Quarterly</u> 68 (1987): 57-70.

[8] Roberts's world-based approach says (via her world-based N*) that in this example the world in which I expose Jones to a high probability of a horrible existence, but in which Jones does not come into existence is a permissible world. Again, I say that this has little normative significance until it is connected with the permissibility of action. And it had better say that my action was wrong. Otherwise it is not appropriately sensitive to probabilities.

[9] For helpful comments I thank Earle Conee Tony Ellis, Brad Hooker, Trenton Merricks, Gene Mills, and Melinda Roberts.