

Week 3: Person-affecting theories, part I

Reading list

Core

Narveson, 'Utilitarianism and new generations' (on Weblearn)

Arrhenius, 'Population ethics: The challenge of future generations', chapter 10 ("presentism, necessitarianism and actualism").

This is an unpublished book manuscript. Arrhenius has given permission for the manuscript to be shared with participants in this class, on the condition that it not be circulated further. Therefore, in particular, I can't put it on Weblearn. I will email it to all those present at the Week 2 class. If you weren't present in Week 2 but plan to attend in Week 3, please email hilary.greaves@philosophy.ox.ac.uk for a copy.

Broome, *Weighing lives*, sections 10.2 ("The intuition of neutrality") and 10.3 ("counterexamples to the principle of equal existence") (available online via SOLO).

Further readings on person-affecting theories in general (relevant to weeks 3, 6 and 7 of this seminar)

There is an enormous amount of literature on person-affecting approaches. What follows is a partial list of things some of which you might like to look at at some point, in no particular order. (Obviously, we don't expect anyone to read all or even most of this!)

If you know of any good "person-affecting" literature that is not in this list, I'd be keen to hear of it, as I am supposed to be writing a survey article on this for *Philosophy Compass* later this year.

Narveson, "Moral Problems of Population" = **26 pages**

Arrhenius, Chapter 9 ("person-affecting theories")

McMahan, "Problems of Population Theory" (33 pages)

Bigelow and Pargetter, "Morality, Potential Persons and Abortion" (10 pages)

Shiffrin, "Wrongful Life, Procreative Responsibility, and the Significance of Harm" (33 pages)

Hanser "Harming Future People" (23 pages)

Parsons, "Axiological Actualism" (11 pages)

Hare, "Voices from Another World: Must We Respect the Interests of People Who Do Not, and Will Never, Exist?" (27 pages)

Wasserman, "Hare on De Dicto Betterness and Prospective Parents" (8 pages)

Howard-Snyder, “Damned if you do; damned if you don’t!” (15 pages)

Roberts, “The Asymmetry: A Solution” = **35 pages**

McMahan, “Asymmetries in the Morality of Causing People to Exist” = **20 pages**

McMahan, “Wrongful life: Paradoxes in the morality of causing people to exist” (35 pages)

Roberts, “Can it Ever Be Better Never to Have Existed At All? Person-Based Consequentialism and a New Repugnant Conclusion” (27 pages)

Roberts, “A New Way of Doing the Best That We Can: Person-based Consequentialism and the Equality Problem” (35 pages)

Roberts, *Child versus Childmaker: Future Persons and Present Duties in Ethics and the Law* (254 pages)

Meacham, “Person-Affecting Views and Saturating Counterpart Relations” (31 pages)

Beckstead, *On the Overwhelming Importance of Shaping the Far Future*, Chapter 4 (20 pages)

Bader, “Neutrality and conditional goodness” = **79 pages**

Bader, “The asymmetry” = **29 pages**

Temkin, “Intransitivity and the Mere Addition Paradox” = **51 pages**

Ross, “Rethinking the Person-Affecting Principle” = **34 pages**

Norcross, “Intransitivity and the Person-Affecting Principle” (8 pages)

Roberts, “Temkin’s essentially comparative view, wrongful life and the mere addition paradox” (21 pages)

Population Ethics

TT15 week 3: Person-affecting theories, part I

Hilary Greaves

1. The basic intuitions of person-affecting approaches to population ethics
 - a. Slogans (Narveson):
 - i. “Morality is about making people happy, not making happy people.”
 - ii. “Morality is about how we treat whatever people there are.”
 - iii. “How large a population you like is purely a matter of taste” (i.e., is not a matter of *morality*).
 - b. Where average utilitarianism (in the *population*-axiology sense) goes wrong: Sure, it’s great if you can raise the average utility/happiness level of a *fixed* group of people. But it doesn’t follow that it’s an improvement to “increase the average happiness of a given group” *by expanding the composition of that group*, so as to include some previously-excluded and happier-than-average people (Narveson’s example of the Fervians).
 - i. And analogously for total utilitarianism.
 - ii. To make mistakes(?) like this, one would have to be thinking of people as ‘mere containers for happiness’.
 1. (The average/total-utilitarian’s likely reply: I didn’t, in fact, generate my population axiology by hasty generalisation from the fixed-population case, via considering people as mere containers or otherwise.)
 - c. Suppose we add an extra person, and that person is happy (has a life worth living), and no-one else’s well-being level is affected. Have we made things better for one person and worse for no-one?
 - i. If so, then it follows from the (pretty uncontroversial) Pareto Principle/Principle of Personal Good that the answer is ‘yes’.
 - ii. Person-affecting intuition (though): No, because *it is not the case that having a life worth living is better for the person who lives it than never existing at all* [cf. week 5].
 1. Although we do need to capture the wrongness of predictably creating lives of unmitigated suffering (“the asymmetry”).
 - iii. Further, being *better-for-at-least-one-person* is a necessary condition of being better-simpliciter. (?)
 - d. We don’t have any moral reason to have (additional) children, even if those children would predictably (i) be happy and (ii) have a positive net effect on the lives of others, and even if it would cost us little or nothing to have those children.
 - i. It’s not just that we don’t have any obligation to have such children (*that* could be accommodated by a deviation from maximising consequentialism).
 - ii. Hard to accommodate the (alleged) datum that we don’t even have any reason, except by denying that the resulting state of affairs is better.
2. The basic challenge for a person-affecting approach to population ethics: *but seriously, what is your theory, in the end?*

- a. A non-exhaustive taxonomy of the possibilities
 - i. Postulate a (complete) axiology that only counts (or counts for more) the interests of people who 'already exist'. (This week)
 - 1. General worry: crazy verdicts on some cases; some versions also have 'structural problems'
 - ii. Postulate incompleteness in the axiology: In at least some cases involving a pair of different/different-sized populations A and B, it is not the case that A is better than B, nor that B is better than A, nor that A and B are equally good. There is a fourth option: A and B are *incomparable/incommensurable* in terms of goodness. (Formally: neither $A \succcurlyeq B$ nor $B \succcurlyeq A$.)
 - 1. General worry: either inconsistency, or too much incomparability. (Week 6)
 - iii. Do without axiology: state principles about what one ought to do *directly*, i.e. without going via any notion of overall goodness.
 - 1. General worries:
 - a. Capture the obvious fact that one ought to bring about Bliss rather than (disjoint-population) Hell/Mediocrity.
 - b. Find such a theory with any significant elegance/internal coherence/systematicity/predictive power to it. (Will not bother everyone (cf. Kamm).)
 - c. Not clear that this helps, anyway (see Arrhenius, ch. 12).
3. 'Impartiality-violating' ways of making the person-affecting idea precise (Arrhenius, chapter 10)
- a. Consider again the idea that the only people whose interests need be taken into account are the people who 'already' exist. What does 'already' mean, here?
 - i. Presentism: the people who exist *now*.
 - 1. Motivation: an explicitly agent-relative morality. One's "moral club" consists of one's contemporaries.
 - ii. Actualism: the people who *actually* exist (sometime, possibly in the future, and possibly as a result of the action we actually choose).
 - 1. Motivation: "There is no basis for the protection of merely possible persons"; merely possible persons don't have any interests; there *are* no merely possible persons.
 - iii. Necessitarianism: the people who will exist *regardless of which option we choose* (from our current set of options).
 - 1. Motivation: morality is about how we treat whatever people there *anyway* are.

4. Disenfranchised people problems (DPPs)

- a. Basic idea: construct a clearly-better outcome A and a clearly-worse outcome B such that, however, both A nor B include only non-X people (or, only non-X people are *affected* by the choice between A and B).
- b. DPP for (Strict) Presentism: Future Bliss or Hell
- c. DPP for (Strict) Actualism: Future Bliss Or Hell, again. (Or: Future Bliss Or Muzak-And-Potatoes – to sidestep the 'asymmetry' move)
- d. DPP for (Strict) Necessitarianism: Suppose that

- i. A consists of the x-people at well-being level 10, and the y-people at well-being level 1.
- ii. B consists of the x-people at well-being level 9, and the y-people also at well-being level 9.
- iii. C consists of the x people only (at any well-being level).
Then Necessitarianism implies that A is better than B (relative to this choice set). But that is implausible.

5. Some structural observations

- a. Necessitarianism is “choice-set-dependent (CSD)”, in the sense that: whether or not $A > B$ can depend on which other alternatives, besides A and B, are also in the choice set.
- b. Actualism is (choice-set-independent but) “choice-dependent”, in the sense that: its ranking of the alternatives, for a fixed choice set, depends on which alternative is (actually) chosen.

6. Actualism and Axiological Invariance

- a. Because of Actualism’s choice-dependence, it violates
 - i. **Axiological Invariance:** The moral value of an action does not depend on whether or not that action is performed.
- b. Violations of Axiological Invariance are problematic because:
 - i. They render the theory useless for deliberative purposes?
 - 1. Reply 1: criterion of the right, not decision procedure
 - 2. Reply 2: Ratificationism
 - ii. They can generate implausible moral dilemmas
 - 1. Of a special kind: in any choice situation, there *is* a right action and the action that is right is one that one *could* have performed; but every action G is such that *if* one performed G, then G would be wrong (“performance dilemmas”).
 - a. Example: Let $A=(x:10)$; $B=(x:11, y:-100)$.

7. Problems for choice-set-dependent (CSD) theories

- a. If ‘the’ ordering is relative to choice set, so that in fact what we have is *many* partial orderings $>_{A,B}$, $>_{A,B,C}$ etc, in general, taking the union $U_s >_s$ will yield an ‘ordering’ that fails to be transitive/asymmetric...
 - i. Necessitarian example: if $A=(x\text{-people at } 10, y\text{-people at } 5)$, $B=(y\text{-people at } 10, z\text{-people at } 5)$, $C=(z\text{-people at } 10, x\text{-people at } 5)$, then, according to Strict Necessitarianism, $B >_{\{A,B\}} A$, $C >_{\{B,C\}} B$, $A >_{\{C,A\}} C$.
- b. ... and this is a problem, because...
 - i. ...It’s a conceptual truth that betterness is transitive
 - 1. Replies:
 - a. Call it “betterness*” if you like
 - b. I don’t have any cyclicity (or whatever) *within* a single choice set; and I never said that the *union* of my various choice-set-relative betterness relations was itself supposed to be a betterness relation
 - ii. ...Someone with cyclic preferences can be money-pumped
 - 1. Reply: Not if they see it coming

- iii. ...We cannot in practice know precisely which range of outcomes is in our choice set, so choice-set-dependence makes a theory inapplicable in real decision-making
 - 1. Reply: use expected utility theory, as usual

Week 5: Comparing existence and non-existence

Reading list

Core

Arrhenius and Rabinowicz, “The value of existence” = **19 pages**

Fleurbaey and Voorhoeve, “On the social and personal value of existence” = **26 pages**

[optional] Bykvist, “The benefits of coming into existence” = **27 pages**

Further readings

Parfit, *Reasons and persons*, pp. 487-90 = **4 pages**

Arrhenius, *Population Ethics: The challenge of future generations*, sections 9.3 and 9.4

Benatar, “Why coming into existence is always a harm” (42 pages)

Brown, “Better never to have been believed” (10 pages) (further debate between Brown and Benatar in 2013; see Brown’s “Reply to Benatar”)

Harman, “Can we harm and benefit in creating?” (25 pages) and “Critical study of [Benatar]” (10 pages)

Heyd, “Procreation and Value: Can Ethics Deal With Futurity Problems?” (19 pages)

Velleman, “Persons in Prospect” (68 pages)

Holtug, “On the Value of Coming into Existence” (24 pages)

Population Ethics, TT15

Week 5: Comparing existence and non-existence

Hilary Greaves

1. Some background
 - a. \succsim : better-simpliciter
 - b. \succsim_X : better-for-X
 - c. Incomparability: not $A \succsim_{[X]} B$, and not $B \succsim_{[X]} A$ either (neither better, nor worse, nor equally good)
2. Comparativism
 - a. Better-for-X and well-being
 - i. The set of all possible well-being levels carries a natural ordering (at least a *partial* ordering).
 - ii. The notions of *betterness-for-X* and *X's well-being level* are connected, at least via
 1. Minimal betterness/well-being connection: If Joe has wellbeing level w_1 in A and well-being level w_2 in B, and $w_1 > w_2$, then A is better for Joe than B.
 2. (This leaves the controversial questions open.)
 - b. Many questions in population ethics would become simpler if we could also locate a given person's *nonexistence* somewhere on this well-being scale/somewhere in this betterness-for-that-person ordering.
 - i. Well-being attributions: zero well-being if you don't exist?
 - ii. Better-for-x statements: better-*for-you* to exist (with positive well-being) than not to exist?
 - c. *Comparativism*: the thesis that statements of the form "A is better/worse/equally good for X than B" can be true even when X does not exist in one of the states of affairs A, B.
 - d. Relevance to population ethics: E.g.
 - i. Reconciles PAP with the obvious data in non-identity cases (e.g. Future Bliss and Hell). (PAP: A is not better-simpliciter than B unless A is better-*for-someone* than B.)
 - ii. Supplies a powerful argument for Mere (and Benign) Addition (via Pareto).
 - iii. The *falsity* of Comparativism would make it easier to defend the claim that there are no moral reasons to create happy people (and/but: harder to defend the claim that there *are* moral reasons *not* to create miserable people).
3. The master argument against comparativism
 - a. Broome: "It cannot ever be true that it is better for a person that she lives than that she should never have lived at all. If it were better for a person that she lives than that she should never have lived at all, then if she had never lived at all, that would be worse for her than if she had lived. But if she had never lived at all, there would have been no her for it to be worse for, so it could not have been worse for her."
 - b. Rendering this a bit more formally:

- i. Notation: Let @ denote the actual world. Let A, B, ... range over possible worlds.
 - ii. **(EC: Existence Commitment)** A sentence of the form "A is better for X than B" is true only if X exists.
 - 1. So, in particular: the sentence "A is better than B"
 - a. Is *actually* true only if X actually exists;
 - b. Is *true in A* (respectively, in B) only if X exists in A (resp. B).
 - c. ("True in A": *would* be true *if A were actual*)
 - iii. (LI: Limited Invariance) If a sentence of the form "A is better for X than B" is true in any world, then this sentence is also true in both A and B.
 - iv. (NC: Non-Comparativism) Unless X exists in both A and B, then "A is better for X than B" is not true. (From P1, P2)
- c. Defence of EC
- i. The sentence "A is better for X than B" asserts that a certain three-place relation, *Better-for*, holds of the triple (A,B,X). But no relation can hold of (A,B,X) if any one of A, B, X fails to exist – in particular, if X fails to exist.
 - ii. "I kicked a dog" can be true only if I exist, and there exists at least one
- d. Defence of LI
- i. Obviously not true for an *arbitrary* sentence.
 - ii. Specifically for betterness claims: Whatever we normally say in defence of axiological-invariance principles? (Cf. week 3.)
 - iii. Note (though) that if EC is also accepted, the claim had better be only *limited* invariance. ("A is better for X than B" can be true only if *necessarily* X exists" would be going too far.)
 - 1. This might make it harder to provide an adequate *motivation* for LI.
4. Against the Master Argument's conclusion: 'Comparativist' intuitions
- a. "Each of us, I think, is lucky to have been born." (Nagel?)
 - b. Consider a *guardian angel*: an agent who decides in the interest of the particular person whose guardian angel she is.
 - i. Surely my merely possible sister's guardian angel would have preferred that she be brought into existence if her life would have been great, and not if it would have been terrible? (Although: some issues re. individuation.)
 - ii. Arrhenius (in some early work?) suggests *analysing* "A is better for X than B" as "X's guardian angel would prefer A to B", as a means of bypassing the worries involved in EC. (Gets the cart before the horse?)
 - iii. Alternative point: the apparent truth of "X's guardian angel would prefer A to B" does function as *strong evidence* that "A is better for X than B" is (must be, the 'master argument' notwithstanding) true.
5. Responses to the master argument
- a. Accept the conclusion: Broome, Bader, Bykvist, (Parfit?), ...
 - b. (Accept Existence Commitment but) reject Limited Invariance: Arrhenius and Rabinowicz
 - i. There's no motivation for LI, anyway.

- ii. Existence Commitment gives an *explanation of why* “A is better for X than B” cannot be true-in-A if X does not exist in A, but no analogous reason for thinking that “A is better for X than B” cannot be true-in-B (if X exists in B).
 - iii. So: if X exists in B (with positive well-being level) and does not exist in A, then B would be better for X than A *if B were actual*, but not *if A were actual*.
 - iv. This requires modifying the PAP:
 - 1. “Subjunctive Weak Pareto [SWP]: if A would (if A obtained) be better than B for all the A-people, and A would (if B obtained) be better than B for all the B-people, then A is better than B.”
 - c. Reject Existence Commitment
 - i. Roberts: a person has a well-being level of zero in a world in which she does not exist. (*Comparativism* then follows from *Minimal betterness/well-being connection*.)
 - 1. Bykvist’s retort: “Well-being presupposes being”.
 - 2. Arrhenius and Rabinowicz’s criticism: Roberts’ *argument* for her claim (anyway) is internally inconsistent.
 - ii. Fleurbaey and Voorhoeve: Let A be world in which X does not exist. Then X does not have any *well-being level* in A; but A can be better-worse for X than some other world B (if X exists in B). (Betterness-for-X comparisons are not reducible to facts about X’s well-being.)
- 6. Damage-limitation? The attempt to make do with goodness/badness, jettisoning betterness/worseness (-for-the-person)
 - a. Suggestion: It is never true that existence is *better*, or *worse*, for the person than non-existence. But existence can be (non-comparatively) *good* or *bad* for a person.
 - b. This is (presumably) supposed to explain why it’s better not to create miserable lives: we avert a thing that *would be bad* for people, even though what we do isn’t *better* for anyone.
 - c. Skepticism about this approach:
 - i. Not clear that a fundamentally *non-comparative* notion of good/bad even makes sense (as opposed to: ‘F is good’ just meaning ‘F is better than the implicit status-quo’).
 - ii. Even if it makes sense, it’s morally irrelevant.
- 7. More on rejecting Existence Commitment
 - a. Note that neither of the above attempts to reject EC *explains what is wrong* with the semantics-based argument for EC. Let’s attempt to do this:
 - b. By modifying the semantics?
 - i. Why *must* the semantics for “A is better for X than B” amount to a three-place relation *Better-for(A,B,X)*? Clearly, semantics needn’t always mirror surface grammar. (Cf. “I did it for the sake of my kids.”)
 - ii. This problem is well-recognised, in particular, in the case of attempts to say things about mere possibilities.
 - 1. Cf. “the problem of negative existentials” (Frege, Russell, etc.)
 - 2. F&V: a merely-possible person may “have the property of” “having a great life in such-and-such counterfactual situation.”

- iii. Project: Defend the rejection of EC by appeal to some less-straightforward account of the semantics for “A is better for X than B”.
 - 1. First attempt: *Better-for-X(A,B)*. Does this involve existential commitment only to A and B, or also to X?
- c. By modifying the metaphysics?
 - i. If either modal realism or necessitarianism is true, then merely possible persons ‘exist’ in the sense that is required for them to appear as relata in the semantics.
 - ii. Further, both of these positions is independently motivated.
 - 1. Modal realism: Lewis, ‘On the plurality of worlds’
 - 2. Necessitarianism: Williamson, ‘Modal logic as metaphysics’
 - iii. Project: Defend the rejection of EC by arguing (ideally) that any account that is able to cope with *the phenomena that motivate* modal realism and/or necessitarianism will inevitably also provide the resources for rejecting EC. (Better than *assuming the truth of* modal realism/necessitarianism, because more robust.)
 - 1. E.g.: Whatever the true metaphysics, we must be allowed to *quantify over* mere possibilia, or to do something that is equivalent.

Week 6: Person-affecting theories, part II

Reading list

Recap

Broome, 'Weighing lives', sections 10.2 ("The intuition of neutrality") and 10.3 ("counterexamples to the principle of equal existence") (available online via SOLO; assigned for week 3, so I'm assuming that most people have read it already).

Core

Broome, 'Weighing lives', chapter 12 ("Indeterminate betterness") (available online via SOLO)

Rabinowicz, "Broome and the intuition of neutrality" (on Weblearn)

Broome. "Reply to Rabinowicz", in *Philosophical Issues* 2009 (on Weblearn)

Further readings

Bader, 'Neutrality and conditional goodness' (on Weblearn)

Heyd...(? TBC)

Broome, 'Weighing lives', chapter 11 (available online via SOLO)

Person-affecting theories, part II: The intuition of neutrality

1. (One version of) the basic intuition
 - a. A natural way of capturing the intuition that there's no moral reason to create an extra person (even if that person would predictably have a good life) is axiological: doing so *would not make the overall state of affairs better* (or worse).
 - i. Strong version (Bader, Heyd): Regardless of the well-being level the extra person would have.
 - ii. Weak version (BBD): Provided the extra person's well-being level would lie within a certain range.
 - b. Two prima facie sensible options:
 - i. Other things equal, the resulting state of affairs would be *equally as good as* the original;
 - ii. Other things equal, the two states of affairs are *incomparable* in terms of goodness(-*simpliciter*).
 - c. Main task this week: investigate the prospects for 'incomparability'-based person-affecting theories.
 - d. Main worry for incomparability-based approaches
 - i. We don't want too much incomparability! E.g., it had better not be the case that *every* pair of states of affairs with non-identical populations is incomparable in terms of goodness – that's no better than saying that such states of affairs are equally good. (Future Bliss and Hell, climate change...)
2. Formalising the neutrality intuition in terms of equally-good
 - a. **Principle of Equal Existence (PEE)**: Let A be any population. Let A' be a population consisting of A plus a single additional person, with the well-being levels of persons in A unaffected by the addition. Then, there is some 'neutral' range of well-being levels such that, if the extra person's well-being is within this range, the two distributions are equally good. (~Broome, WL, p. 146)
 - b. Broome's argument against PEE
 - i. Let w_1, w_2 be well-being levels that are both inside the neutral range, but such that $w_2 > w_1$. Let A be any population. Let $A_1 := A + [w_1]$, and $A_2 = A + [w_2]$ (in the obvious notation). Then:
 - ii. By PEE, $A \sim A_1$, and $A \sim A_2$.
 - iii. By transitivity of equally-as-good-as, $A_1 \sim A_2$.
 - iv. But, by the Principle of Personal Good (i.e., roughly, the Pareto Principle), $A_2 > A_1$. This contradicts our conclusion that $A_1 \sim A_2$.
 - c. Possible reply: "the goodness of an alternative is essentially comparative"...
 - d. Broome's taxonomy of (other) replies:
 - i. Deny transitivity of equally-as-good-as (hopeless?)
 - ii. Appeal to 'conditional goodness' (not clear that this generates a distinct option)
 - iii. Say that goodness is relative to population (? Cf. Broome's discussion of Dasgupta, in WL ch. 11.)
 - iv. Say that goodness is indeterminate/appeal to incomparability (TBC...)

- v. Look for a non-axiological way of capturing the neutrality intuition (Roberts?)
- vi. Give up the neutrality intuition
- 3. Formalising the neutrality intuition in terms of indeterminacy/incomparability
 - a. A brief and partial taxonomy of incomparability-like notions
 - i. A and B are *incomparable*: No ‘positive value relations’ (usually taken to amount to: better-than, worse-than, equally-as-good-as) hold between A and B.
 - ii. A and B are *on a par*: A is not better than B, nor worse than B, nor equally as good as B, *nor* incomparable to B. There is a fourth positive value relation of *parity*: to say that A and B are *on a par* is (approx..) to say that they are *roughly equally as good as* one another.¹
 - iii. It’s *vague* whether or not A is better than B: ‘A is better than B’ is like ‘Tom is bald’, if Tom is a borderline case of baldness. (Then plug in your favourite account of vagueness, e.g. supervaluationism/epistemicism/etc.)
 - iv. Indeterminacy, noncomparability, incommensurability...
- 4. Formalising the neutrality intuition in terms of incomparability
 - a. ‘Principle of incomparable existence’ (PIE): as for PEE, but replace ‘equally good’ with ‘incomparable’.
 - b. Replacing PEE with PIE blocks Broome’s argument against PEE, because ‘incomparable to’ is clearly *not* transitive.
 - c. The problem of ‘greedy neutrality’
 - i. Example: Consider the following 4 populations:
 - A= (4, 4, ..., 4, 6, Ω)
 - B= (4, 4, ..., 4, 6, 1)
 - C= (4, 4, ..., 4, 4, 4)
 - D= (4, 4, ..., 4, 4, Ω)
 - ii. By PIE, A, B are incomparable, and C, D are incomparable.
 - iii. But presumably $C > B$ (greater average and total well-being, fixed population, more equal).
 - iv. By transitivity of betterness, we cannot have $A > C$. Broome: this is a ‘greediness’ of incomparability that was not intuitively supposed to be part of neutrality. (And beware an ‘incomparability’ version of the RC.)
 - 1. [Comment: Important to distinguish ‘this was not intuitively supposed to be part of neutrality’ from ‘this is implausible’.]
- 5. Digression: Broome’s positive view
 - a. There is a *single* ‘neutral level’ (not a neutral *range*, as in the neutrality intuition).
 - b. But it’s a matter of *vagueness* which that level is.
 - c. Adopt a supervaluationist account of vagueness. Then,
 - i. We have a ‘representor’ class of CLU value functions, differing from one another on the value of the ‘critical level’.

¹ It’s controversial whether or not this notion of ‘parity’ makes sense (and, if it does, how). See e.g. Chang, ‘Introduction’ to her book *Incommensurability, Incomparability and Practical Reason*; Chang, ‘The possibility of parity’; Chang, ‘Parity, Interval Value, and Choice’; Gert, ‘Value and Parity’; Peterson, ‘Parity, Clumpiness, and Rational Choice’; Rabinowicz, ‘Value relations revisited’.

- ii. A is *determinately* better (resp. *determinately* worse) than B iff A is better than B according to *all* value functions in the representor.
 - d. On the 'Repugnant' and 'Sadistic' Conclusions
 - i. General dilemma for (single-neutral-level) CLU theories: if the critical level is too low (extreme case: if it's zero), then the RC bites much as in total utilitarianism. But as soon as the critical level is at all positive, the theory implies the 'sadistic conclusion'; and the higher the critical level is, the worse the SC-like implications of the theory.
 - ii. General strategy for a 'vagueness' version of CLU: let the range of admissible candidates for the neutral well-being level be $[0, \alpha]$, where α is sufficiently high that the remaining Repugnant Conclusion doesn't seem too bad, and extending the range all the way down to 0 is supposed to in some sense avoid the SC.
 - 1. On the RC: Assuming that the 'lives barely worth living' in the 'Z-world' are within the range of admissible candidates for the neutral level, this theory will deem it a *matter of vagueness* whether the RC is true.
 - 2. Similarly, on the SC: Assuming that the 'added lives worth living' are within the range of admissible candidates for the neutral level, this theory will similarly deem it a *matter of vagueness* whether the positive-wellbeing-lives-added scenario is worse than the negative-wellbeing-lives-added scenario.
 - 3. (Are these consequences OK? Not clear that this is really any better than straightforward incomparability, as above...)
- 6. 'Strong'-neutrality theories
 - a. Heyd's view:
 - i. Any two states of affairs involving *non-identical* populations are incomparable w.r.t. betterness. Different-people choices (including, but not only, different-number choices) lie outside the domain of ethics.
 - ii. Obvious concern:
 - 1. This is precisely the sort of rampant incomparability we wanted to avoid; and/or,
 - 2. We still need to see what the 'non-ethical' part of total theory says about the different-people choices. Designating that part of our total theory "not part of ethics" won't make it any easier to devise a sane theory.
 - b. Bader's view
 - i. The general betterness relation:
 - 1. For fixed population sizes, coincides with total/average utilitarianism.
 - 2. Any two states of affairs involving *different-sized* populations are incomparable.
 - ii. But there is also a 'meta-betterness' relation, which does sometimes deliver positive comparisons between different-sized populations.

1. On one version: a smaller population S is meta-better than a larger population L iff: all |S|-sized subsets of L have lower total/average well-being than S. (This condition is met if e.g. all members of L have a well-being level that is lower than the average in S.)
 - a. And, presumably: L is meta-better than S iff : all |S|-sized subsets of L have *higher* total/average well-being than S.
2. In general, this meta-betterness relation is not transitive. But
 - a. For comparing populations both of which have perfect within-population equality of well-being, the meta-betterness relation coincides with the average-utilitarian ordering.
 - b. In any case, the *transitive closure* of this (non-transitive) meta-betterness relation is the average-utilitarian ordering.
- iii. The *preferability* relation is a combination of the betterness and meta-betterness relations:
 1. In any given choice situation, identify the option or options that involve the smallest population.
 2. No preferability relations hold between these 'minimal' options and the non-minimal ones.
 3. Within the minimal set, the preferability relation follows total/average utilitarianism (since the general betterness relation does).
 4. Within the non-minimal set, A is preferable to B iff either A is generally-better than B (for same-size-population pairs), or A is meta-better than B (for different-size-population pairs).
- iv. This view has a complex verdict on the RC: e.g. it is permissible to choose Z over A. But if the option-set includes all members of a sequence A, B, ..., Z, it is only permissible to choose A (the 'no-expansion' option) or B (the highest-average-WB option from among the 'expansion' options). (Having ruled out A, though, it's then permissible to 'move' from B to C, and so on.)
- v. Some concerns about the view
 1. Has this view just patched together aspects of average utilitarianism and total-incomparability theories, in such a way as to generate no real improvement over the 'average satisfactoriness' of those theories?
 2. Like all choice-set-dependent theories, the theory relies on there being a hard fact of the matter as to what the 'choice set' is. Is the choice set really so determinate?
 3. Threat of dynamic inconsistency