

LEAP

Law, Ethics and Philosophy

Law, Ethics and Philosophy (LEAP)

LEAP is a new peer-reviewed international journal dedicated to work in ethics, legal theory, and social and political philosophy. It welcomes clear, rigorous and original submissions that address concrete issues of public concern as well as more abstract theoretical questions. It also has the distinctive aims of (a) fostering work drawing on a variety of disciplines within the social and natural sciences, including biology, economics, history, and psychology; and (b) promoting dialogue between the Anglophone and non-Anglophone worlds.

We invite submissions of articles up to 10,000 words, discussion notes up to 5,000 words, and replies and exchanges not exceeding 3,000 words. All published submissions will be subject to blind review, and the journal will notify authors of submitted pieces about the progress of their submission within six weeks.

Although LEAP accepts exclusively submissions in English, the journal strongly encourages submissions from authors who also write in languages other than English, and will always strive to ensure that their work is assessed on the basis of its content and not primarily its mode of expression. Where necessary the editorial process provides authors with guidance regarding matters of English style. The journal is published digitally by Marcial Pons (www.leap_journal.com), the leading publisher of analytical legal philosophy in the Spanish-speaking world.

Enquiries regarding the journal may be directed to: leap.editors@marcialpons.es.

LEAP
Law, Ethics and Philosophy

Vol. 2
2014

Marcial Pons

MADRID | BARCELONA | BUENOS AIRES | SÃO PAULO
2014

Editorial Board

Editors

José Luis Martí, *Pompeu Fabra University*
Hugo Seleme, *National University of Córdoba,*
Argentina
Paula Casal, *ICREA-Pompeu Fabra University*

Andrew Williams, *ICREA-Pompeu Fabra*
University
Serena Olsaretti, *ICREA-Pompeu Fabra*
University

Associate Editors

Ingvild Almås, *Norwegian School of Economics*
Samantha Besson, *Fribourg University*
Jordi Ferrer, *University of Girona*
Ernesto Garzón Valdés, *Johannes Gutenberg*
University, Mainz
Cristina Lafont, *Northwestern University*

Genoveva Martí, *ICREA-University of Barcelona*
Lukas Meyer, *University of Graz*
José Juan Moreso, *Pompeu Fabra University*
Félix Ovejero, *University of Barcelona*
Zofia Stemplowska, *University of Oxford*

Editorial Board

Aulis Aarnio, *Tampere University*
Lucy Allais, *University of the Witwatersrand*
Elizabeth Anderson, *University of Michigan*
Richard Arneson, *University of California, San*
Diego
Gustaf Arrhenius, *Stockholm University*
Michael Baumann, *University of Düsseldorf*
Juan Carlos Bayón, *Autonomous University,*
Madrid
Carmen Bevia, *Autonomous University,*
Barcelona
David Bilchitz, *South African Institute for*
Advanced Constitutional, Public, Human
Rights, and International Law
Geoffrey Brennan, *University of North Carolina*
at Chapel Hill
Ian Carter, *University of Pavia*
Joseph Chan, *University of Hong Kong*
Thomas Christiano, *University of Arizona*
Bruno Celano, *University of Palermo*
Antony Duff, *University of Minnesota*
John Ferejohn, *New York University*
Víctor Ferreres, *Pompeu Fabra University*
Roberto Gargarella, *University of Buenos Aires*
Robert Goodin, *Australian National University*
Axel Gosseries, *University of Louvain*
Lori Gruen, *Wesleyan University*
Riccardo Guastini, *Genova University*
Alon Harel, *Hebrew University of Jerusalem*

Daniel Hausman, *University of Wisconsin-*
Madison
János Kis, *Central European University*
Matthew Kramer, *University of Cambridge*
David Lefkowitz, *University of Richmond*
Kasper Lippert-Rasmussen, *Aarhus University*
Frank Lovett, *Washington University in St. Louis*
Stephen Macedo, *Princeton University*
Jeff McMahan, *Rutgers University*
Jane Mansbridge, *Harvard University*
Adèle Mercier, *Queens University, Ontario*
Liam Murphy, *New York University*
Ingmar Persson, *University of Gothenburg*
Philip Pettit, *Princeton University*
Thomas Pogge, *Yale University*
Wlodek Rabinowicz, *Lund University*
Joseph Raz, *Columbia University*
Debra Satz, *Stanford University*
Julian Savulescu, *University of Oxford*
Seana Shiffrin, *University of California, Los*
Angeles
Anna Stilz, *Princeton University*
Victor Tadros, *University of Warwick*
Larry Temkin, *Rutgers University*
Jeffrey Tulis, *University of Texas at Austin*
Philippe Van Parijs, *University of Louvain*
Georgia Warnke, *University of California,*
Riverside
Ruth Zimmerling, *Johannes Gutenberg*
University Mainz

Editorial Assistants

Antonio Barboza, Leticia Morales, José Ezequiel Páez, Jahel Queralt, Lorena Ramírez, Laura Roth
Pompeu Fabra University
www.leap-journal.com

CONTENTS

Law, Ethics and Philosophy (LEAP), Vol. 2, 2014

	<u>Page</u>
1. What Do We Owe to Poor Families?	9
RICHARD J. ARNESON	
2. Firms, States, and Democracy: A Qualified Defense of the Parallel Case Argument	35
ÍÑIGO GONZÁLEZ-RICOY	
3. Rethinking the Good - A Small Taste	61
LARRY TEMKIN	
4. In Defense of the Internal Aspects View: Person-Affecting Reasons, Spectrum Arguments and Inconsistent Intuitions	91
ÓSCAR HORTA	
5. On the Asymmetry of “Better than”, Person-Affecting Views, Spectrum Arguments, the “Unhelpfulness” of Rejecting Transitivity, and Implications for Moral Realism	113
LARRY TEMKIN	
6. Internal or External Grounds for the Nontransitivity of “Better/Worse than”	125
INGMAR PERSSON	
7. Intransitivity and the Internal Aspects View	145
LARRY TEMKIN	
8. Globalization and Global Justice in Review	159
NICOLE HASSOUN	

	<i>Page</i>
9. Libertarian Welfare Rights: Can We Expel Them?	173
CHARLES GOODMAN	
10. Social Contract Theory in the Global Context	185
PETER STONE	
11. Sufficiency, Equality and the Consequences of Global Coercion...	199
KOK-CHOR TAN	
12. Legitimate Coercion: What Consent can and Cannot do.....	219
NICOLE HASSOUN	
13. Two pictures of the global-justice debate: A reply to Tan.....	229
LAURA VALENTINI	
14. A Brief Rejoinder to Valentini.....	237
KOK-CHOR TAN	

What Do We Owe to Poor Families?

RICHARD J. ARNESON

University of California, San Diego

Abstract

This essay argues that when there is a moral duty to procreate, nonprocreators owe assistance in the task of providing for children, even if their presence renders nonprocreators worse off. When new children bring benefits to nonprocreators, they have a duty of reciprocity owed to cooperating parents. If there is a moral duty to provide meaningful work opportunities, especially to the worse off, we have special duties to help poor people enjoy opportunities for the meaningful work of raising children. Given the benefits of stable families for both their adult and child members, justice requires facilitating the enjoyment of stable family life by poor people.

Keywords: Procreative duties, reciprocity, meaningful work, marriage promotion, social justice.

1. INTRODUCTION¹

Women in economically advanced societies and in some other locales have gained increased opportunities to participate in the market economy and public life in the past century.² These gains are undeniably great advances in social justice. They have been accompanied by strains. It's unlikely that so far we have anywhere evolved the right mix of institutional adjustments and changes in social norms and practices to facilitate the transition to a world in which men and women contribute on equal terms in the labor market and the entrepreneurial arena.

In this essay I try to offer a perspective on what we owe to poor families in the context of recent social changes, especially women's increasing participation in economic life outside the home.³ "Poor families" refers to adults and children living in a household with low income and wealth, roughly

1. For helpful comments and criticism I thank Serena Olsaretti and Andrew Williams.

2. Fuchs, 1988; also Esping-Andersen, 2009.

3. Although my discussion stays at an abstract level above social policy choice, I suspect my thinking tends to focus on U.S. problems and circumstances. On the differences between

the bottom quintile. “We” refers to the nonpoor.⁴ Of course what we owe to people we fundamentally owe to individual persons not groups; focus on families is a heuristic for public policy guidance.

Offering a perspective is here a distinct and separate enterprise from assessing candidate fundamental moral principles. The aim is to identify appealing mid-level norms that might gain wide allegiance among people who differ in their fundamental moral allegiances. Offering a perspective is also distinct and separate from advancing public policy proposals. In order to be in a position to advance and defend a specific public policy proposal, one needs to be able to show that if implemented in our actual circumstances the policy would lead to outcomes that are morally desirable (without violating moral constraints). Doing that requires a comprehensive empirical understanding of relevant actual circumstances to which this essay does not aspire.

My procedure is to suggest how to think about how to fulfill our obligations to people in disadvantaged families on the assumption that what we owe to others by way of cooperating on fair terms or lending a helping hand depends on what policies would do to help people live genuinely better lives, have richer and more fulfilled lives rather than bleak or squalid ones.

John Rawls once wrote that his proposed theory of justice, justice as fairness, does not look behind the uses that people make of their resources and opportunities in order to measure, much less maximize, the satisfactions they gain (Rawls, 1993). Provided basic institutions are arranged so that the distribution of resources and opportunities turns out to be fair, what individuals do with their resources and what quality of life they fashion for themselves is their business, not the business of society. So Rawls urges. In contrast, I assume that determining what policies would be fair requires us always to be looking past the distribution of liberties and opportunities to see what impact the policies are having on the quality of the lives of the individuals who are affected. (The fact that’s what just and fair depends on what’s good in this way is fully compatible with paying attention to personal responsibility in the all-things-considered determination of morally desirable policy.)

welfare state policy in the U.S. and in Europe, see Garfinkel *et al.*, 2010; also Alesina and Glaezer, 2004.

4. For the most part I confine my attention to relations between poor and nonpoor in a single political society. This scope restriction sets aside duties that people have to poor people anywhere, and at any time, regardless of who inhabits what society. This is just a simplification for purposes of discussion; I don’t in fact believe people within a single political society owe more to fellow members than to outsiders.

I do not attempt to come up with a complete set of norms regulating what we owe to poor families. I urge that when a poor person becomes the parent of a child and is willing to assume childrearing responsibilities for that child, the rest of us acquire strong obligations to help bring it about that the child starts adult life well socialized and with good prospects. In one range of cases, the new child expectedly depletes available resources and hence her arrival renders the rest of us worse off. Nonetheless, the persons responsible for bringing the child into existence may be acting to fulfill a moral obligation that falls collectively on each generation to “be fruitful, and multiply” and that places duties on each member individually to contribute a fair share to the project of procreation and childrearing and adequate provisioning. The extent of what is owed depends on the amount of burden that childbearing and childrearing place on procreators and in part on the amount of burden that helping more or helping less would place on others. In another range of cases, an additional moral obligation falls on nonprocreators. When the addition of a new child to the world expectedly leads to increase of wealth and culture and tends to make those living in the vicinity of the new child better off, the project of childbearing and childrearing is morally comparable to a beneficial cooperative practice and those who benefit from the operation of the practice are duty bound not to free ride on the cooperative efforts of others but rather to contribute their fair share of the burdens of the practice.

An additional consideration that generates a duty on the part of the non-poor to aid the childrearing efforts of poor parents is that for people whose labor market prospects are poor, the opportunity to raise children is a very significant, perhaps the only feasible opportunity they have to engage in creative and fulfilling work. In this situation, assisting people to undertake and successfully complete a parenting project may be a requirement of distributive justice owed to them.

Finally, and tentatively, I suggest that people have a defeasible moral right to stable nurturing family arrangements. Children have a right to a decent home environment. Adults have rights to freedom to date and mate on mutually agreeable terms, and a right to a social environment that facilitates successful steady family arrangements. Such a social environment will educate youth to be disposed to seek long-term steady family arrangements and to have the skills and personal traits needed for success in this venture. These family-oriented rights and obligations do not negate the entitlements to fair treatment of those who will avoid family entanglements as adults, but these entitlements properly construed are consistent with society’s implementing policies and promoting norms that nudge individuals toward stable nurturing family arrangements.

2. DUTIES TO PROCREATORS

When a poor person has a child, what, if anything, is owed to the childbearer in virtue of this event? Let's back up. When any person has a child, what, if anything, is owed to the childbearer in virtue of this event?

A robust libertarianism holds that the new child has a right to decent prospects in life, but that the duty to supply these decent prospects falls only on the individuals responsible for bringing this child into existence and on no one else. A variant of this view holds that the duty to care for a child falls in the first instance only on these responsible individuals, and duties fall on others only as a second-best back-up responsibility that comes into play only if the responsible procreators fail to do what they ought. This is a possible view, but here I set it aside without comment (Steiner and Vallentyne, 2009).

Suppose instead that we all have a duty to do our fair share to help provide each new child with decent life prospects. That means that nonprocreators have a duty to contribute to fair shares for children under conditions of full compliance (when procreators are contributing what they ought to give). But we might wonder what is a fair division of this burden across the procreators of the child and everyone else. After all, in the standard case, those who bring a child into existence either engage in sex with the aim of producing a child or engage in sex with the understanding that a possible outcome of what they are doing is that a new child might be brought into existence. So maybe the procreators bear some special responsibility in this regard.

This issue is insightfully analyzed in a resource egalitarian framework by Paula Casal and Andrew Williams (Casal and Williams 2004). Other philosophers inspired by the resource egalitarianism of Ronald Dworkin adopt similar views.⁵ They contrast two possible cases. In one case, bringing a child into existence makes others better off, by increasing the supply of resources available for humans to use. In another possible case, bringing a child into existence makes others worse off, by decreasing the supply of resources available for humans to use. One can discern an asymmetry between the cases. When people voluntarily choose to have children, to the point that there is no undersupply, with resulting benefits to others, we who benefit from the parents' childbearing and childrearing efforts do not owe compensation to the parents for this benefit. This is a positive externality and those who benefit from it do not thereby incur any obligation towards those who produce it to reward them for doing so. In contrast, when people voluntarily choose to have children, and thereby make others worse off, there is a moral case for requiring the responsible procreators to pay the costs their childbearing generates and not seek to impose these costs on nonprocreators.

5. See, for example, Rakowski, 1993; also Clayton, 2006.

To illustrate, suppose society begins with a group of adults forming a society with a supply of unowned resources available for fair distribution among them. No one has prior claims on the resources. The resources should then be divided fairly among the individuals. Following the resource egalitarian views of Ronald Dworkin, Casal and Williams suppose a fair distribution is the one that mimics the outcome of an equal auction in which all resources are put up for bid and the individuals are given equal bidding power (equal money for use in the auction) and there is trade to equilibrium, with the added proviso that there are also simultaneously in play hypothetical insurance markets for handicaps and native marketable talents.⁶ Individuals as they bid for resources are also able to purchase insurance against suffering handicaps, with the overall incidence of handicaps known but not the particular risk that one has one or several, and able to purchase insurance against having low marketable talent. In this hypothetical market one knows one's native talents but not what prices they are likely to fetch when the auction ends and economic life commences. The insurance pays out if one has the covered condition and one pays into the fund that makes these payments if one lacks the condition against which one has insured. To simplify, assume the outcome of the hypothetical auction and insurance markets is that all of the individuals receive an equal share of available material resources.

Now suppose some people voluntarily act so that a new child is brought into the world, and suppose the existence of this child will lessen not increase the available stock of resources. Suppose that as each new child enters the world, each is owed an equal resource share as defined by the hypothetical auction and insurance market procedures, and just suppose the outcome continues to be that each new person should get an equal share of available resources. The existing adults must together then be worse off, must accept fewer resources, to satisfy the just claim of the new child. Who should bear this cost? Casal and Williams point out in effect that if we start from a fair initial distribution and there is a fair framework for interaction after that (roughly, a standard private ownership free market economy with the requirement that one not harm others without their consent), then those and only those who have voluntarily brought about the child should pay for the costs the child's entry into the world imposes on others.

From the perspective of the nonprocreators, the cost of the new child created by others is bad brute luck, luck that falls on them beyond their power to control.⁷ This luck merits full compensation. In contrast, the cost of the new child in its relation to the child's voluntary creators is option luck,

6. Dworkin, 2000, chapters 1 and 2; also Dworkin, 2011, chapter 16. See also Steiner, 1995.

7. The claim in the text here is advanced by Rakowski, not by Casal and Williams. On the distinction between brute luck and option luck, see Dworkin, 2000. See also Vallentyne, 2002.

costs brought about that they should have foreseen and might have avoided. There is a case then for requiring the procreators to absorb the costs of the new child, including the costs of giving the child resources to enable her to have fair initial prospects should fall on the procreators and no one else. In this respect having children is like building a fire on your property for your own purposes that emits pollution that fouls the air that others must absorb. These costs imposed by your voluntary conduct should be borne by you and no one else, so you owe full compensation to those your behavior would otherwise be harming, in the absence of this full compensation. As in the pollution case, we suppose that it is morally permissible for agents to act with their resources for their own purposes in ways that have spillover negative effects on others (unless there are special circumstances such as that the negative effects are noncompensable) if and only if they fully compensate others for any damages incurred.

It bears emphasis that Casal and Williams are assuming background conditions of fair distribution of resources. Their analysis and assessment would not straightforwardly apply to a world like ours in which the distribution of resources over time fails to conform to the resource egalitarian justice principles. So their analysis and assessment does not straightforwardly yield any implications for what we might owe voluntary procreators who bring costly children into existence and what they might owe us when social relations are already marred by distributive injustice.

There is some plausibility to the account that Casal and Williams provide. However, the view they offer is incomplete and thus defective. Let us look at the two cases they consider.

2.1. Case one: procreation imposes costs on nonprocreators

To see the difficulty, imagine a world with very low population. This might be the situation in a world shortly after the events of the Adam and Eve story as told in the Judeo-Christian Bible have unfolded. Or we might imagine a world with very reduced population in the wake of some natural or man-made disaster such as a war that wipes out almost the entire human population and sets us the task of starting human society afresh. Suppose there are four individuals and resources are initially divided fairly in line with the hypothetical equal auction and insurance markets. Again, to simplify, just suppose the fair distribution is an equal distribution. Each individual gets one-quarter of the Earth's material resources. The four individuals then proceed to save and consume and build with the resources they own. They trade with each other on mutually agreed terms. All is well, as assessed from the resource egalitarian perspective.

Now imagine that two of the individuals pair up and have two children. This addition of new people might over time reduce or increase the resources available for human use; let us suppose there is a reduction. In these circumstances, the new individuals have a right to a fair initial share of resources; let us suppose this is a share of resources equal to what each of the four initial persons received. Here the resource egalitarian position yields the clear result that the two procreators and they alone should bear the cost of introducing the new people into the world and providing them fair initial shares.

This result seems clearly mistaken. Or at least, an issue needs to be faced, to settle who owes whom what in this setting. In effect Rakowski's assessment of the situation assumes that the resource endowments that the four people initially get are lifetime entitlements come what may. The four are in effect lords and ladies of the Earth, entitled to all of it, fairly divided. There would be no moral impropriety if the four all lived out their lives as nonprocreators and the total population of the Earth after the start of our account turns out to be four. Each might say, "One-fourth of the Earth is mine, to use as I choose."

But someone might protest that each person's initial endowment of resources only provisionally belongs to him. Each is a partial steward of the Earth, with a responsibility to pass it on intact, or perhaps to pass along some combination of material resources plus technology so that future generations of people get fair shares and decent life prospects. Moreover, there is a responsibility of some sort bearing on each of us to bring it about the future generations exist. There is a moral duty falling on the four initial owners to bring about future people.

I would add, there is a duty not simply to maintain current population but, as the biblical injunction says, to "be fruitful, and multiply."⁸ Exactly what the correct secular version of the biblical injunction requires would involve elaborating a full population ethics, which I am not able to do. But even without having in hand a full population ethics, we can see some of its contours, and can say with assurance that when the Earth can sustain increasing population with good lives for people, it is not morally permissible for existing people to decline to reproduce and let the human race die out. Nor for that matter would it be morally permissible merely to sustain a very low population, as in our toy example. Here I am appealing to a vague but controversial premise, which those who oppose the claim that nonparents owe help to those who voluntarily create costly children may reject. However, rejection comes at a cost. If the resources of the Earth are abundant, or can predictably support a very large population for the indefinite future given predict-

8. The Book of Genesis 1: 28.

able improvements in technology that render natural resources increasingly useful, many will agree with me that it is wrong for us to fail to produce the population increase that can bring huge gains in lives worth living.⁹

Of course there are some ways in which the addition of new people to the world might worsen the prospects of already existing people that would plausibly be regarded as trampling on their genuine rights. Suppose the story of how adding new people worsens the existing people's life prospects is that we procreators culpably mismanage the education and socialization of our children and they grow up to be antisocial, disposed to engage in acts that wrongfully harm others. If this is the way that the already existing people's life prospects are made worse off by the addition of new people, the already existing nonprocreators may have a plausible case that the procreators have done them wrong. (I say "may have a case", putting the claim tentatively, because we have yet to see what responsibilities nonprocreators might have regarding the education and socialization of new people.)

But simply being made worse off by the arrival of new people on Earth because one has to share the Earth with them, given that their arrival was beyond one's power to control, and brought about by the actions of others, does not introduce a justice claim of nonprocreators against responsible procreators. To see whether any such entitlement of nonprocreators is violated, we need to look at the duties and obligations that we all have with respect to childbearing.

I have suggested that there is a collective duty to be fruitful and multiply, a duty that falls on all of us together to bring about sufficient population growth or maintenance (or reduction, in unfortunate circumstances). This collective duty generates individual duties, but in a conditional and indirect way. Consider by way of analogy the duty that falls on an enormous crowd of people lying around at the beach, to carry out a rescue when someone falls in peril of drowning. There is a duty initially that falls on each of us to carry out a rescue if no one else does so, a duty that disappears when someone able to carry out the rescue commences it. The collective duty then is transmuted into a duty to provide help to the rescue team if that is needed, and to compensate for the costs they incur and the services they render, and to participate in follow-up efforts to help the imperiled person recover from the near-drowning incident either by contributing labor directly or by contributing to a fund of resources used for these amelioration efforts.

We need not enter into the abstract question, what theory of morality best explains and justifies the particular shape that the duty to rescue that binds us has. Suffice it to say that according to any moral theory that stands a chance of being right, morality contains a significant beneficence requirement—a

9. For a view contrary to what is asserted in the text, see Narveson, 1967.

requirement to make the world better by one's efforts. The beneficence requirement may be multifaceted, but it is plausible to insist that it includes a requirement to contribute to making the world better by bringing the number of people who enjoy good life prospects closer to what it should ideally be.

In the context of the initial situation of four people living on Earth, a resource egalitarian view might initially assign each of us provisional ownership of one-quarter of the Earth's material resources. But this is not a permanent bequeathable property right over the entirety of those resources. Nor is it a full property right that disappears only with one's death, as the example we are considering illustrates. The initial distribution of resources takes place against a moral background in which population growth is mandatory and there is an obligation to share resources with new people whether or not the addition of the new people worsens one's situation compared to what it would have been had no new people arrived on the scene. That is not a morally relevant baseline of comparison because one never has any right to enjoy throughout one's life an undisturbed ownership of the equal share of resources that is implemented at an earlier time on the ground that it is fair to then existing individual persons.

So in my example the procreators are fulfilling a collective duty binding on all existing people, and nonprocreators are obligated to contribute to this mission to some degree, even if population increase takes away from their initial endowment of resources, where the initial allocation of these resources is assumed to be (provisionally) fair. At a minimum the nonprocreators are obligated to accept this reduction in their resources that we are supposing accompanies the morally mandatory population increase.

Notice that the resource reduction accompanying population increase that according to Casal and Williams triggers a duty on the part of procreators to make good this loss to nonprocreators (or prevent it from ever occurring by absorbing themselves the costs of giving their children fair initial shares) need not even involve any worsening of the lives of the nonprocreators all things considered. The presence of the new humans might be pleasurable for everyone to contemplate. Watching children frolic is fun. Hence the population increase might leave no one sad that this event has occurred, without this fact counting against the Casal-Williams claim of the nonprocreators to full compensation for resource losses. This feature of their view is generated by its being resource-oriented, not welfare-oriented.

The position I am sketching regarding procreation obligations need not deny that special responsibilities fall on the particular persons who voluntarily act to produce childbirth. These persons have brought about the existence of particular needy and helpless human infants at a particular time, and surely doing so triggers a special duty of care for the welfare of the

dependent beings one has created. Seeing this is compatible with placing voluntary childbearing in a broader context in which we all have duties to contribute to population increase.

What holds true in a four-person world can also hold true in a world already populated with billions of people. Again, I don't presume to be in possession of a satisfactory population ethics principle or suite of principles.¹⁰ Any of a range of principles will yield the plausible implication that the population of the Earth, given present circumstances, ought to increase. Also, suppose that the right population ethics calls for population stabilization not increase in our circumstances. Carrying out this mandate might still lead to the circumstance that triggers the duty of procreators to absorb costs of procreation and pass none of them along to nonprocreators according to Casal and Williams. Again, those who bring about births necessary for stabilization are still fulfilling a collective duty, and their acts trigger duties falling on nonprocreators, on the view this essay proposes.¹¹

The claim then is that the voluntary procreator whose childbearing and childrearing incurs costs we all must share is relevantly unlike the polluter who acts for his own purposes in ways that impose spillover harms on others. He is more like the voluntary rescuer when a person in peril is threatened with drowning and many persons might come to the rescue. His act helps to fulfill a duty we all owe and we all have duties to help carry through the fulfillment.¹² Either the rescue effort itself, or the cost of reimbursing those who bear special risk or expense, might impose costs on nonrescuers, but these are costs that nonrescuers are morally obligated to bear, up to some point. Same goes for nonprocreators.

2.2. Case two: procreation brings gains to nonprocreators

Turn now to case two. Suppose some people voluntarily chose to have children—again we are supposing this occurs against the backdrop of an

10. For a useful introduction to population ethics issues, see Broome, 2004.

11. I do not deny that in some circumstances adding people to the world will be morally wrong according to a reasonable population ethics. In this case procreators might be wronging nonprocreators by imposing unfair costs on them. But whether this is so depends on the shape of the collective duty we all share, in some circumstances, to bring about population decrease. It might turn out that nonprocreators have not done all that they ought to do, to facilitate population decrease, so the mere fact that Alf has been directly responsible for adding a child to the world, when this outcome is morally undesirable, does not by itself establish that Alf has wronged nonprocreating bystanders.

12. I assume that Casal and Williams would agree that if there is a duty to procreate that falls in one way or another on all of us, that duty might generate obligations on the part of nonprocreators to assist in the childrearing and resource provision for children that voluntary procreators are morally required to give their children. They will disagree that there is any such duty.

initially fair distribution of resources— and nonprocreators gain spillover benefits. In resource egalitarian terms, which we are here not challenging, the addition of new people to the world brings it about that there are more resources for everybody. In another terminology, bringing children into the world, in some circumstances, generates positive externalities that fall on others, including nonprocreators. In this scenario, do nonprocreators have some duty to assist with costs of childbearing and childraising that is triggered by their receipt of these external benefits, on the assumption that they neither asked for nor consented to the imposition of these benefits?

This question calls to mind the Hart-Rawls principle of fairness, which holds that “when a number of persons engage in a mutually advantageous cooperative venture according to rules, and thus restrict their liberty in ways necessary to yield advantages to all, those who have submitted to these restrictions have a right to similar acquiescence on the part of those who have benefited from their submission.”¹³ The idea is that when procreation and childrearing are correctly regarded as a cooperative practice that fits this characterization, the cooperating behavior of the participants generates duties of reciprocity and fair play that fall on those who are recipients of benefits. As Rawls remarks, “We are not to gain from the cooperative labors of others without doing our fair share” (Rawls, 1999: 96).

Casal and Williams do not deny that obligations can arise from the Hart-Rawls principle of fairness. But they interpret the principle in a way that restricts its applicability. In this connection we might consider doubts about how obligations might arise under Hart-Rawls raised in separate discussions by Robert Nozick and A. John Simmons.¹⁴ Roughly, the idea is that those who incur obligations to cooperators under the Hart-Rawls principle must either voluntarily accept the benefits of the scheme or at least be willing to accept benefits voluntarily if voluntary acceptance were possible. Mere receipt of benefits does not suffice to obligate. Also, those to whom duties are owed under Hart-Rawls must be intending to benefit others by their cooperative actions undertaken under a fair scheme of rules. Merely acting in ways that happen to benefit others does not suffice to generate obligations of reciprocity to repay. Along a similar line, Casal and Williams say “the principle concerns nonexcludable goods that are produced by cooperative activity in which individuals bear some cost, which they would not otherwise bear, in order to produce the good”. (Casal and Williams, 2004: 159).

13. Cited from Nozick, 1974: 90.

14. See Nozick, 1974: 95; also Simmons, 1979, chapter 5; also Simmons's contribution to Simmons and Wellman, 1999, for defense of the principle of fairness, see Arneson, 1992; Arneson, 2013.

Casal and Williams anyway hold that in our world, procreation and childrearing as actually practiced do not meet the conditions of the Hart-Rawls principle of fairness suitably interpreted. Hence, procreative activities do not give rise to obligations on the part of nonprocreators to bear a share of the costs of childrearing and fair provisioning of children with resources as they become responsible adult citizens.

In response, the first point to note is that if there is a collective duty to expand or sustain population and individual duties flow from that collective duty, then in the case in which the childbearing and childrearing activities of procreators confer external benefits on others, the requirement falling on recipients of these benefits to do more to assist in the child production enterprise in virtue of this receipt of benefits holds whether or not the conditions of the Hart-Rawls principle of fairness apply. Returning to the analogy with rescue cases, suppose some people undertake a rescue in circumstances in which all of us, including me, bear some responsibility for undertaking rescue and sharing its costs. Suppose the rescue effort happens to shower benefits on some bystanders, who are themselves obligated to share the costs of the rescue. The receipt of benefits alters what qualifies as the fair cost sharing arrangement. If some in the group of those who are snared in the web of collective obligation to bring about rescue happen to gain side effect benefits and others similarly obligated do not, those who benefit should pay a greater share of the overall cost of the rescue operation and those who do not so benefit should pay less.

Nonetheless, I hold that the childbearing and childrearing enterprise in many communities does qualify as a cooperative practice satisfying the conditions of the Hart-Rawls principle and so triggering obligations of reciprocity to contribute one's fair share.

First point: if other conditions are met, mere receipt of benefits can sometimes suffice to trigger obligations. Some goods provided via cooperative schemes are nonoptional with respect to a group of people: if anyone in the group consumes any benefits, everyone must consume some benefits.¹⁵ When nonoptional benefits are delivered by cooperation, one need not vol-

15. There are further distinctions that need to be drawn here. Sometimes if any members of a group consume any of a good, all must consume the same amount of it. Here the

good just falls on people and there is no choice whether or not to consume it. Sometimes all must consume some if any group members consume some, but a different amount of the good might fall on different group members—again, all of this happening in the absence of anyone's voluntary choice as to how much to consume. In still other cases even though a good is nonoptional for members of a group, how much of the good each individual consumes is up to that individual. In this last case, the extent of one's obligation to help pay for the good may depend on how much one chooses to consume ("may depend", because consumption may be nonrival, and one's choice to consume may be just a decision to avoid

untarily accept benefits in order to become obligated to pay one's fair share. Nor need one's will be disposed to voluntary acceptance if that were possible. Casal and Williams cite Ronald Dworkin as criticizing versions of the principle of fair play that assume that "people can incur obligations simply by receiving what they do not seek and would reject if they had the chance." Dworkin comments, "This seems unreasonable."¹⁶ Casal and Williams do not definitely embrace whatever amendment of the principle Dworkin's claim seems to suggest. I would definitely deny that we should endorse the Dworkin criticism as stated in the first instance. Suppose a cooperative practice is operating that provides national defense for the community you inhabit. Suppose you have false empirical beliefs, and do not see that the national defense practice really does provide you genuine benefits, worth their cost, in a scheme whose costs are fairly apportioned. Your subjective disbelief that you benefit does not obviate your obligation to contribute under the scheme. The same holds, I would say, if your error is moral rather than factual. Suppose you have misguided pacifist views and believe benefiting by threat of violence under any circumstance is wrong. Nonetheless, the threat of violence that maintenance of national defense involves is in fact morally right and you do in fact benefit, and benefit from a morally acceptable practice that is fair. I say, in these circumstances, you have an obligation to contribute, indeed an enforceable obligation, which your subjective opinion to the contrary does not obviate.

Suppose that the production-of-children practice does in fact confer benefits on me, and the benefits are worth the cost, and the costs are fairly apportioned by the rules of the practice. Suppose also that at least some of the benefits provided are nonoptional. Just by living in the society, I cannot avoid receipt of benefits. These circumstances do not yet suffice to establish that obligations arise under Hart-Rawls. The enterprise must be a cooperative venture, fairly organized, and the cooperators must be intending by their activity to be conferring the benefits of the scheme on others. Also, the cooperators must be incurring costs under the arrangements. The objection then insists that these further conditions are not satisfied.

There is an interesting question raised here, which this essay will not seek to answer. The question is what obligations if any arise in situations in which some but not all of the Hart-Rawls conditions are met. For example, suppose that the conditions of the Hart-Rawls principle of fairness are satisfied, except that the "cooperators" don't think of themselves as acting to benefit others, but if the fact that their efforts do spread benefits widely through the

waste, in the context of a scheme that is not worth its costs and should cease to exist). These complications do not affect any claim I urge in this essay.

16. Dworkin, 1986:194 (cited after Casal and Williams, 2004: 168).

community were brought to their attention, they would find this welcome news, and acquire an intention so to benefit the others. We might call these people latent cooperators. Do their latently cooperative activities generate obligations in those who benefit from their efforts to pay a share of their costs? I suspect the answer is “Yes” but will not pursue the issue here. I simply contend that the plain unvarnished conditions of the Hart-Rawls principle, rightly understood, are standardly satisfied by participants in child production practices, so given that these practices shower benefits on others in the neighborhood, the obligations of reciprocity that the fair play principle generates here are triggered.

One issue is whether participants in the practice are acting to benefit others. I grant that people who either have sex with the aim of having children or who have sex and then make a decision to bring the fetus to term when it is discovered that a pregnancy has started normally act in the expectation that raising children will enrich their lives and make the decision for this reason. But there is normally another element in play. People decide to have children for self-fulfillment, but this is a moralized notion of self-fulfillment. Procreators think that their childrearing activities will significantly enhance the community in which they live, and they are also aware, perhaps in a somewhat inchoate or vague way, that there is a duty to be fruitful and multiply that falls on their community and is one that their procreative choices help to fulfill. People’s motives are mixed, but that does not preclude their having the motivations that are conditions for Hart-Rawls to apply. After all, many who volunteer to contribute to national defense, the paradigm of a cooperative scheme to which Hart-Rawls applies, have mixed motives and aim in part at their own self-fulfillment, through meaningful work or glory seeking or the like.

Here is a relevant comparison. Take the standard example of a public goods provision scheme that generates duties under the principle of fairness. Bandits periodically menace peaceful farmers living near each other in a narrow valley. Some farmers initiate a protection system. Valley dwellers are to take turns standing sentry duty each night, which will reduce the losses of all to predatory bandits. Suppose some people really hate standing sentry duty, some don’t mind, some fancy the activity even though it is risky. So a fair and tolerably efficient scheme for distributing the burdens of the protection scheme might involve asking for volunteers, and then if the number of individuals who volunteer is adequate, requiring other valley inhabitants to pay into a fund that compensates the volunteer sentries for their noble activity and provides special health care benefits to sentries injured on duty, etc. The sheer fact that people volunteer to supply the needed public good (partly for altruistic, partly for self-interested motives) does not negate the moral appeal of the claim that others who benefit from their activities on behalf of the community owe them compensation.

Someone might object that if people voluntarily act in ways that benefit themselves and spill benefits also on others, without any offer of compensation, no compensation to them is owed—and this is the situation of procreators vis-à-vis benefiting others. In reply: Under the Hart-Rawls principle of fairness reasonably interpreted, obligations can arise on the part of those who benefit from cooperative schemes that shower nonoptional benefits on a group of people (or for that matter on those who voluntarily seek and get optional benefits from such a public goods delivery system), *even if the cooperators are net beneficiaries from the scheme in the absence of contributions from nonparticipant beneficiaries*. The cooperators who gain on balance in this way can still be unfairly treated by the free riders, and the scheme can be made more fair in its distribution of benefits and burdens if nonparticipant beneficiaries are required to pay a fair share. (What constitutes “fair shares” is a topic not addressed in this essay.)

I grant that procreators normally expect their lives to be improved, in prudential terms, by having and raising children. Nonetheless, procreators also expect to bear some costs they would prefer to avoid, or lessen, if they could, despite their expectation of overall gain. And procreators recognize that in deciding to have children they are inevitably making a risky choice, that if things go badly might result in large-scale disruption or degradation of their lives. When children turn out badly, even through no fault or oversight of the parents, the parents’ lives can be blighted, pretty much destroyed. I submit that these generally applicable characterizations of the decision to have and raise children suffice to satisfy the condition of sacrifice incurred by cooperators that is required for the Hart-Rawls principle to apply.

Readers may wonder what benefits unavoidably fall on nonprocreators arising from the childrearing efforts of parents in their society. These vary. Some are highly local. Only those living nearby get the benefit of seeing your children gamboling along the street. Some are widely diffused. All people benefit, as they age, from reinvigoration of the culture as a result of the creativity of youth, and of stimulus to the economy from the energy and ambition of the young.¹⁷

17. In an interesting essay, Serena Olsaretti raises doubts as to whether obligations to share the costs of having children arise under Hart-Rawls along the lines I have urged. My discussion in the text answers these doubts. She makes a further suggestion: that as modern societies are actually organized, significant benefits of parenting are coercively channeled to nonparents, as when children are taxed to pay for old age assistance programs and other elements of the modern welfare state. This sounds plausible. However, I worry that critics might respond that in an ideally just society, these arrangements would be made on a pay as you go basis, and there would not be forcible takings of the sort she highlights. For example, each generation could save for its own future, and use the proceeds to purchase goods needed in old age. See Olsaretti, 2013.

One might raise another worry. The Hart-Rawls principle of fairness stipulates that obligations arise, given certain conditions, when people participate together in a mutually cooperative venture according to rules and others receive benefits from the scheme. Is the production-of-children practice sensibly regarded as carried out according to rules? Yes. In a just society, laws and social norms stipulate what those who contribute to and benefit from childbearing and childrearing owe to one another. There are recognized expectations. In a state of nature, the rules are more inchoate, and essentially consist of the recognized duty to procreate according to population ethics along with the principle of fairness and a sensible range of interpretations of the duties and obligations thereby generated.

The preceding scrappy remarks have tried to support the idea that non-procreators have obligations to procreators in a wide range of circumstances. Your bringing about the birth of a child generates obligations in bystanders to help in the upbringing of the child.

3. DUTIES TO PROVIDE OPPORTUNITIES TO THE POOR TO BE SUCCESSFUL PARENTS

Whatever we owe to procreators and children generally, we owe more to procreators who are poor and to children born into poverty.¹⁸ To return again to the rescue analogy: if there is a drowning threat on a beach filled with people, a few people undertake the necessary rescue of those who are in peril, and the rescuers then turn out to be especially vulnerable and incur large costs in the course of the rescue, or turn out to need extra help in order to complete the rescue effort successfully, the bystanders have an extra obligation to compensate the rescuers who have sustained great loss and especially to assist these would-be rescuers in bring their rescue efforts to success.

A wide variety of approaches to social justice will converge on the judgment that we owe more to worse off (as compared to better off) members of society who are engaged in childbearing and childrearing. Under a wide range of circumstances, utilitarian, egalitarian, sufficientarian, and prioritarian views will affirm this judgment.

I want to focus on the subset of poor children and poor adult family members who are also low in the ensemble of personal traits that constitute native ability.¹⁹ Some poor children and their guardians are below average

18. See Bou-Habib, 2012.

19. For an interesting defense of the view that in determining people's fundamental political status and political entitlements a duty of opacity respect is owed all agents above a thresh-

in bank account wealth and income but blessed with high levels of talent. Some are cursed with low talent endowments. (Ability is multifaceted and multidimensional, but I suppose some are poorly endowed all things considered; for simplicity let's just speak of those with low talent.) With luck and pluck the impoverished people with low ability may end up leading great lives, but they are surely likely to be clustered among the people who end up badly off in overall lifetime well-being.

A further point is perhaps worth noting. Our duties to poor people likely include a duty to provide them a fair opportunity to become parents and successfully raise children. This duty applies with special force to the subset of poor individuals with low marketable skills.

Parenting is meaningful and creative work. (Of course, it is also drudgery and frustration; that does not undermine the claim just stated.) The parent has the duty to form the child's character, to shape the lumpish infant into a particular person. This is in some respects a creative, artistic task, like painting a picture or making a sculpture, with the special twist that from the very beginning the material on which one is doing the creating is a conscious human with agency interests of its own, which gradually come to include interests in self-making and self-determination. It is as though you were trying to paint a work of art on a canvas that had arms and hands of its own that were clasping paint brushes, and the canvas arms were making their own strokes and painting over your brush strokes—not to mention that there are many other forces besides your artistic efforts that are shaping the outcome of the process, including some deliberate painting by other people's hands. Moreover, the canvas arms gradually become more adept and eventually take over the process, if you are successful.

For many poor people, and especially for those who suffer the double burden of being born into below-average wealth and being born with a weak endowment of native talent potential, raising a child will be their best option for meaningful and creative work. This will be brutally true under circumstances of injustice, but we should entertain the possibility that a potentially tolerably just society might generate a distribution of income and wealth and a distribution of labor market opportunities that give some individuals meager prospects. If such a society is to qualify as tolerably just, the opportunities of those with below-average prospects must be good enough, and surely must include opportunities for meaningful and creative tasks. There will then be a special obligation to encourage people to undertake childrearing by making sure they have reasonable prospects of being successful parents if they do choose this life path.

old of rational agency capacity, and that opacity respect requires us to refrain from assessing or measuring individuals' agency capacities, see Carter, 2012.

This line of thought is strengthened by the further consideration that parenting requires a different set of capacities than most kinds of creative and challenging work that positions in the market economy provide, especially those that are available to those with low marketable skills at the bottom of the economic heap. A caring, loving disposition, attentiveness and persistence, and some common sense will go a long way toward making one a good parent, especially in the crucial early years of a child's life. (Agencies other than parents can help foster autonomy and sophisticated skills in older children.) Parenting tends to be an accessible form of meaningful work. So a just society that provides adequate opportunities for meaningful work for all its members will be careful to provide, within the mix of meaningful work options made available to poor individuals, good opportunities to be successful parents.²⁰

Much that should be done to help poor (that is, nonaffluent) parents to be successful as parents involves familiar welfare-state measures. We should channel public funds for public education to ensure high quality primary and secondary education along with avenues to higher education for children of poor parents. We should provide high-quality day-care that is available to children of the working poor, to help parents combine paid employment and parenting. We should regulate labor markets to bring about decent employment prospects for individuals with small income and wealth and weak marketable skills, if necessary setting up the state as decent employer of last resort (see Arneson 1990). We should set income tax policy so that individuals with below-average marketable skills are rewarded for seeking and sustaining paid employment and paying self-employment. We should facilitate adult education that develops improved parenting skills. I suggest that in addition the state should be encouraging people to form marriage (or other long-term stable relationships) that make parenting easier by sharing its burdens and for that matter encouraging people, especially men, to regard sustained commitment to parenting as a valuable life option.

4. MARRIAGE PROMOTION, STATE NEUTRALITY, AND MARRIAGE ABOLITIONISM

One might first of all object that the state simply has no business interfering in the decisions of sane adults regarding living alone, cohabitating, marrying, or divorcing. The proper functions of the state do not include regulation of people's friendships and romantic lives. The state has a proper in-

20. I don't deny that a just society in some circumstances might restrict rights to have children and provide fair opportunities for meaningful work to its members in other ways. For example, consider scenarios of severe overpopulation and a drastic need to reduce the birth rate.

terest in securing an adequate upbringing for all children, but this job can and should be done without wrongly interfering in people's private lives. In a diverse society marked by reasonable pluralism of belief, people will differ in their views on sex, romance, marriage, living alone versus living cohabiting versus living communally with several adults, and so on.²¹ So the state should be strictly neutral on this broad issue, neither promoting marriage nor discouraging it.²²

This is in some ways an appealing stance, but notice that if you adopt it, you would seem to be committed to opposing the contemporary movement in some contemporary societies (for example, the U.S.) to establish the legal right for same-sex couples to marry and have the legal privileges of marriage that are restricted to opposite-sex couples in many jurisdictions. (You could still accept as a demand of non-ideal justice that if there is legal establishment of marriage, the status ought to be equally open to couples of both orientations, without being committed to legal establishment of marriage.) The normatively compelling position would be to abolish the legal status of marriage as a state-protected legal status for anyone, of whatever sexual orientation.

The counterargument against marriage abolitionism is that a long-term committed romantic relationship, combining friendship, sex, and the building of a life together, is a valuable achievement, and the official recognition of this type of commitment by the state is a reasonable, noncoercive encouragement to people to attempt this achievement. The goods to be had thereby can equally well be obtained in same-sex and opposite-sex relationships, and the achievement toward which the marital status aspires is equally valuable no matter what the sexual orientation of those attaining it. Hence there should be no discrimination between same-sex and opposite-sex relationships in the state's policy of marriage recognition and endorsement, and this recognition and endorsement are preferable to a hands-off neutral policy, if the state's policy of putting a thumb on the scale in favor of marriage is likely to bring it about that people are better off and the gains of the gainers are not brought about unfairly at the expense of any losers. The point here is that to defend same-sex marriage rights rather than abolition of state recognition of marriage, one needs to argue that long-term relationships are instrumentally and noninstrumentally good for people and that society should foster what is good for people. In a slogan, one needs to oppose state neutrality on the good. I suggest we ought not be marriage abolitionists.

(Although the discussion above refers to same-sex and opposite sex *couples*, nothing hangs on the fact that the group of adults seeking recognition

21. The invocation of reasonable pluralism of belief is claimed to support state neutrality on the good. See Rawls, 1996; also Quong, 2010.

22. This view is vigorously advanced in Brake, 2012.

of their pledge to long-term commitment consists of exactly two people. Larger groups might well form similar bonds and seek the same type of recognition.)

The condition that the state should recognize and endorse marriage on the ground that people gain overall only if gains and losses are not unfairly divided is nontrivial. Some people will do better living alone, or living with their parents or other close relatives, or cohabiting in shifting groups. Is favoring marriage disfavoring them? In many countries in recent years, your chances of getting married and staying married vary with your income. The poor do it less and stick with it less than the nonpoor. Heaping subsidies on marriage might be unfair to the poor.

Here a comparison might be made to recreational drug prohibitions. Suppose the state bans the recreational consumption of methamphetamine, cocaine, and heroin on paternalistic grounds. Such a policy must confront the fact that some citizens are no doubt hurt not helped by the legal ban. Ideally there should be different legal policies for people who would benefit from having these drugs available and for those who would be harmed, but such a legal regime may be unfeasible. If gainers gain enough and losers lose little enough, the ban may be roughly fair, and understanding this, all should abide by it. Much the same might be true of state policies encouraging romantic involvement and in particular long-term committed romantic involvement. The policies might be roughly fair on balance despite the fact that they produce winners and losers.

A significant residual difficulty is that hard drug prohibitions might well impose especially hard burdens on the poor who violate the prohibitions and are then more likely to be exposed to onerous legal penalties than non-poor violators. In the same way, marriage promotion policies will benefit some and hurt others, but among those who end up disfavored, those who are poor will be more likely than nonpoor to be grievously afflicted. If those disfavored in these ways will tend to be the worse off among the worse off, difficult tradeoff issues arise. A policy that helps the worse off may be bad for the worse off segment of the worse off. Depending on the costs and benefits and the numbers of people involved, a morally sensitive cost and benefit calculation might sometimes yield the judgment that an instance of this sort of hard tradeoff is acceptable all things considered. But surely in our public policy choices we should seek ways of avoiding kicking those who are already down.

In considering the desirability from a social justice standpoint of pro-natalist and pro-marriage public policies, we need to be considering not simply what is desirable from the standpoint of an adult person who is poor, but also what is good for poor children. This section has suggested that pro-

moting stable romantic commitment might be fair on balance in its effects on adults even if the policies produce some winners and some losers. The consideration of children's interests complicates the picture.

5. MARRIAGE PROMOTION AND CHILDREARING ASSISTANCE

In several contemporary societies, single-parent households are increasing, especially among the poor. There is evidence that children are likely to fare better in stable two-parent (or two-guardian) households, and that poverty exacerbates whatever problems for children growing up in a single-parent household involves. There is also some evidence that growing up in a single-parent household, with a female as lone parent, creates more difficulties for boys than for girls (Autor and Wasserman 2013). So perhaps the state ought to be discouraging childbearing and childrearing among people, mostly women, who are likely to end up in single-parent households? I suggest the answer is No, but the issues are tricky, even murky.

When two individuals not involved in a stable relationship produce a child, it is not axiomatic that marrying the man who got you pregnant is a good idea, for you or the child. When a marriage is conflictual, it is not obvious that staying together and fighting is better for the children in the household than separating and divorcing. Working class single-parent mothers in the U.S. are more likely than their European counterparts to marry, but they are also more likely to divorce and cohabit again and perhaps remarry again. As a sociologist studying U.S. marital patterns has put it, we have in the U.S. a “marriage-go-round.” Moreover, working class women in the U.S. are more likely than their European counterparts to form nonlasting cohabitation arrangements—the partner-go-round. For children, the evidence is that unstable household arrangements during early childhood years are stressful for children and interfere with healthy development.

How should public policy tilt? Providing financial incentives to marry might exacerbate the marriage-go-round. Providing financial incentives that encourage stable marriage might pressure women to stay in abusive relationships.

One might be tempted by the thought that policies that discourage poor women from having children when they cannot reasonably foresee a stable household arrangement with two or more dedicated adults playing parental roles for their children, if they were successful, would reduce the incidence of single-parent households, a desirable outcome for children.²³

23. This is a theme of political conservatives in the U.S. See for example the collected works of Charles Murray.

If poor women are less prone than their wealthier counterparts to form stable companionate partnerships, and policies enacted to discourage single-parenting would deter some from having children, that outcome should strike us as harsh. For a very large number of prime-age adults, being successful at childrearing is a great part of their good. For many adults, the job of childrearing is the most interesting, rewarding, challenging, and creative work they have the opportunity to perform. As discussed in section two of this essay, this is more likely to be the case for poor adults, whose labor market options are meager.

There are better alternatives to discouraging childrearing on the part of an adult who would be living alone. We should be trying to help poor individuals who want to be parents to succeed in this role rather than to discourage them from undertaking it.²⁴ This is compatible with encouraging stable cohabitation. Also, since men can benefit from successful parenting as well as women, we should be seeking sensible policies that encourage poor men to want to take on childrearing roles and help them succeed in these tasks. There is an element of “the hat makes the man” here. If I take on a responsible role, I am more likely to become committed to it and personally identified with it, and more likely to function as a responsible role-player. We need to seek policies that will induce the man to put on the hat.

The weakening of marriage among working-class people in current times is not mysterious. Compared to earlier times, in our time the benefits of marriage for working-class adults are less and the costs of not getting married are less. For starters, men are less marriageable. They offer less in economic security terms, and women now have other options for securing their economic well-being.

A familiar feature of contemporary economic life is that economies are hollowing out, with fewer good jobs for the uneducated and increasing rewards to those with higher education credentials at the level of college degrees and beyond. In the U.S., “by 1996, the average thirty-year old husband with a high school degree earned 20 percent less than a comparable man in 1979” (Cherlin, 2010: 163). Increasing assortative mating in marriage patterns increases the strains on poor and near-poor people seeking mates. “The winners in the new economy are marrying each other and

24. There is a tradeoff here. Individuals vary in their capacities for good parenting, and vary in how propitious their circumstances are for this social role, and some individuals (rich and poor!) should be discouraged from assuming a childrearing role, even under ideally supportive social conditions. This is consistent with maintaining that generally speaking, people, especially the poor, should be encouraged both to take on the parental role and to form stable family-type relationships, for their sakes and for the sake of the children they will raise, and that those among the poor who have weak parental skills should be helped and supported so that their parenting ventures are successful.

consolidating their gains.’ (Cherlin, 2010: 179; see also Esping-Anderson, 2009: 59-70).

We should assume that these trends will continue for the foreseeable future. In that case, what strategies make sense for poor adults starting out in life? What preparation and—to the extent this might be effective—guidance might society provide them to cope with the world they will face, and what forms of help will facilitate their succeeding in their life aims?

Men with a high school education or less will not offer particularly attractive economic prospects to prospective mates. If these men are predominantly oriented to seeking short-term gratification with male buddies, they won't appear, or be, good candidates for the role of loving partner in intimate relationships. A feminization of their socialization would render them more companionable, more cohabitable, and more marriageable. Even if my job prospects are bleak, if I have serious nonwork interests that I care about and that make me interesting, if I want long-term friendship on egalitarian terms with someone with whom I will also build an intimate stable romantic relationship, and if my life plans are open to the possibility of committing time and energy into childrearing, I look more credible in the eyes of someone seeking a stable romantic partner. Labor force attachment helps as well. If it is just not in the cards that I will be a good traditional breadwinner, at least I can be a stable, rather than erratically intermittent, crumb-winner.

In interviews, high school educated U.S. women report that they have the same life goals as their more educated and wealthy sisters. In a potential mate they seek economic security, intimate friendship, and cooperative faithfulness. Perhaps the “realistic utopian” aspiration for high school educated women would be to find the latter two of the three, and be happy with that. Of course, economic insecurity tends to erode one's capacities for intimacy and loyalty. (I assume that in a just society the distribution of income and wealth can be unequal, but the human cost of economic insecurity that falls on the poor surely affects the degree to which justice requires distributive arrangements that eliminate or cushion that insecurity.)

I have been speaking of people seeking heterosexual partnerships, but there are other sorts. If women's sexuality is more labile than men's, women potentially have the option of choosing to orient sexually toward women and seek female long-term romantic mates and childrearing partners. Or at least, they have the option to the degree that their culture does not wrongfully rule out this option as socially taboo. (I assume men are more likely to be stuck with whatever sexual orientation is planted in their genes.)

If more stable partnership arrangements would be good for people, the simplest way in which social arrangements can facilitate successful partnerships among the poor is by providing generous financial assistance to peo-

ple who are making serious efforts to fashion a good life for themselves. That probably would involve income supplements conditional on labor force attachment. Such income assistance would ease the ever-present grinding strain of hovering on the edge of poverty or being engulfed in it that wears down people's attempts to make lasting loving relationships.

Society could also be generous, rather than stingy, in the educational opportunities provided to children of less educated and economically marginal parents. Better teachers, a longer school day, high quality preschool instruction for very young children, a longer school year, flexible after-school child care, all targeted at poor children, would help the recipients and also help their parents fulfill their important life goal of being successful parents (Waldfogel, 2006). And completing the circle, education of all children, not only the children of the poor, should include a didactic character-forming component. Think of life skills classes that essentially amount to cognitive behavioral therapy—here are the circumstances you can expect to encounter in your future life, here are the difficulties and challenges you will face, what goals do you have and anticipate having and what personal skills and traits will you need to achieve them? This is what I have in mind under the heading of the feminization of socialization, insofar as we are considering the schooling of men.

Regarding state support to the parenting efforts of poor parents, a delicate dance is needed. On the one hand, society should help parents succeed at parenting, for their sake and for the sake of their children. On the other hand, some nonwealthy parents (some wealthy parents too, but that points to other issues) are ineffective parents, and channeling aid to children entirely through their incompetent parents is not an effective strategy for helping children at risk.²⁵ Yet helping poor children in ways that bypass their parents can undermine these parents' authority and self-confidence and decrease their ability to steer their children's lives for the better. To some extent one wants to structure aid to children so that from a very early age they are exposed to authority figures (other than parents) who are visibly working for their benefit and worthy of their trust and emulation. In pre-school and school and day-care centers, though children's peers will probably make a larger imprint on their socialization than their teachers, good teachers can be part-time substitute parents and significant role models. Same goes for the nurse or social worker who makes regular state-sponsored visits to troubled family homes. There is no contradiction in pursuing goals that, in actual circumstances, partially conflict. One seeks to advance the ensemble of the worthy pertinent policy goals appropriately weighted so that one can discern which to pursue more vigorously and to what degree in cases of conflict.

25. This is one lesson that might be drawn from Mayer, 1997.

6. CONCLUSION

In this essay I argue that given population ethics obligations that fall on all of us collectively, we owe assistance to procreators in providing a fair start in life to their children, even on the assumption of an initially fair distribution of resources. We also owe people a fair opportunity to be successful parents. I suggest a perspective for determining what shape and form of public policies these obligations press us to endorse. The perspective assumes that what we owe one another depends on what is intrinsically good, what constitutes a better as opposed to a worse quality of life for the individual living it. I support the perspective largely by illustrating how it would structure deliberation of public policy.

BIBLIOGRAPHY

- Alesina, A., and Graezer, E. L., 2004: *Fighting Poverty in the U.S. and Europe: A World of Difference*, Oxford: Oxford University Press.
- Arneson, R., 1982: "The Principle of Fairness and Free-Rider Problems". *Ethics* 92: 616-633.
- 1990: "Is Work Special? Justice and the Distribution of Employment", *American Political Science Review* 84: 1127-1147.
- 2013: "Paternalism and the Principle of Fairness", in *Paternalism: Theory and Practice*, ed. D. Gauthier and M. Williams, 134-156, Cambridge: Cambridge University Press.
- Autor, D., and L. Hoxby, 2013: *Wayward Sons: The Emerging Gender Gap in Labor Markets and Education*, Washington, D.C.: *Third Way*.
- The Book of Genesis.
- Bou-Habib, P., 2012: "Parental Subsidies: The Argument from Insurance", *Politics, Philosophy, and Economics* 12: 197-216.
- Brake, E., 2012: *Minimizing Marriage: Marriage, Morality, and the Law*, Oxford: Oxford University Press.
- Broome, J., 2004: *Weighing Lives*, Oxford: Oxford University Press.
- Carter, I., 2004: "Respect as the Basis of Equality", *Ethics* 121: 538-571.
- Casal, P., and M. Williams, 2004: "Equality of Resources and Procreative Justice", in *Dworkin and His Critics*, ed. J. Burley, 150-169, Oxford: Blackwell.
- Cherlin, A. J., 2010: *The Marriage-Go-Round: The State of Marriage and the Family in Contemporary America*, New York: Vintage Books.
- Clayton, M., 2006: *Justice and Legitimacy in Upbringing*, Oxford: Oxford University Press.
- Dworkin, R., 1986: *Law's Empire*, Cambridge: Harvard University Press.
- 2000: *Sovereign Virtue: The Theory and Practice of Equality*, Cambridge: Harvard University Press.
- 2011: *Justice for Hedgehogs*, Cambridge: Harvard University Press.
- Esping-Andersen, G., 2009: *The Incomplete Revolution: Adapting to Women's New Roles*, Cambridge: Polity Press.
- Fuchs, V. R., 1988: *Women's Quest for Economic Equality*, Cambridge: Harvard University Press.

- Garfinkel, I., Rainwater, L., and [Smeeding, J.](#), 2010: *Wealth and Welfare States: Is America a Laggard or a Leader*, Oxford: Oxford University Press.
- [Meyer, S. E.](#), 1997: *What Money Can't Buy: Family Income and Children's Life Changes*, Cambridge: Harvard University Press.
- [Narveson, J.](#), 1967: "Utilitarianism and New generations", *Mind* 76: 62-72.
- [Nozick, R.](#), 1974: *Anarchy, State, and Utopia*, New York: Basic Books.
- [Olsaretti, S.](#), 2013: "Children as Public Goods?", *Philosophy and Public Affairs* 41: 226-258.
- [Quong, J.](#), 2010: *Liberalism without Perfection*, Oxford: Oxford University Press.
- [Rakowski, R.](#), 1993: *Equal Justice*, Oxford: Oxford University Press.
- [Rawls, J.](#), 1993: *Political Liberalism*, New York: Columbia University Press.
- 1999: *A Theory of Justice* (revised edition), Cambridge: Harvard University Press.
- [Simmons, A. J.](#), 1979: *Moral Principles and Political Obligation*, Princeton: Princeton University Press.
- [Simmons, A.](#), [Harsanyi, J.](#), [Wellman, C.](#), 2005: *Is There a Duty To Obey the Law?* Cambridge: Cambridge University Press.
- [Steiner, H.](#), 1980: *Essays on Rights*, Oxford: Blackwell.
- [Steiner, H.](#), and [Vallentyne, P.](#), 2009: "Libertarian Theories of Intergenerational Justice", in *Intergenerational Justice*, ed. A. Gosseries and L. Meyer, 50-76, Oxford: Oxford University Press.
- [Vallentyne, P.](#), 2002: "Equality, Brute Luck, and Initial Opportunities", *Ethics* 112: 529-557.
- [Waldfofel, J.](#), 2006: *What Children Need*, Cambridge: Harvard University Press.

Firms, States, and Democracy: A Qualified Defense of the Parallel Case Argument*

ÍÑIGO GONZÁLEZ-RICOY

Universitat de Barcelona

Abstract

The paper discusses the structure, applications, and plausibility of the much-used parallel-case argument for workplace democracy. The argument rests on an analogy between firms and states according to which the justification of democracy in the state implies its justification in the workplace. The contribution of the paper is threefold. First, the argument is illustrated by applying it to two usual objections to workplace democracy, namely, that employees lack the expertise required to run a firm and that only capital suppliers should have a say over the governance of the firm. Second, the structure of the argument is unfolded. Third, two salient similarities between firms and states regarding their internal and external effects and the standing of their members are addressed in order to assess the potential and limits of the argument, as well as three relevant differences regarding the voluntariness of their membership, the narrowness of their goals, and the stiffness of the competition they face. After considering these similarities and differences, the paper contends that the parallel-case argument provides a sound reason in favor of democracy in the workplace—a reason, however, that needs to be importantly qualified and that is only *pro tanto*.

Keywords: firms, states, parallel-case argument, workplace democracy.

* Previous versions of the paper were presented at ICADE and at the Chaire Hoover d'éthique économique et sociale at the Université de Louvain. I am grateful to these audiences, and to audiences in Barcelona, London, Madrid and Princeton, for the central concern of the paper was persistently raised by them. For comments, I am also grateful to Borja Barragué, Jahel Queralt, two anonymous referees for this journal and especially to Andrew Williams. The usual disclaimer applies. Funding was provided by the ARC project 09/14-018 on 'Sustainability' (French speaking community of Belgium) and the Catalan Agency for Management of University and Research Grants.

1. INTRODUCTION

Over the last decades democracy has rapidly expanded worldwide. While in 1946 only 20 out of 71 independent states were democratic, the number increased to 48 in 1989, to 77 in 1994, and to 92 in 2009 (Marshall and Cole 2009: 10-11). The expansion of democracy in the state, however, has not gone hand-in-hand with a parallel expansion of democracy in other realms, such as the workplace.¹ Rather the contrary. The number of cooperatives may have decreased over the last decades.² And attempts to extend the German system of co-determination have either been blocked—as in the case of the Fifth Directive drafted by the European Community in the 1970s—or failed altogether—as in the case of the Bullock Report in the UK (Gold, 2005; 2010). Further, once a battle cry among workers and a central research topic in industrial relations, labor economics, and political philosophy (Christie 1984: 112-128), since the 1980s workplace democracy has attracted a declining interest among scholars, political parties, and workers alike.

Yet, the recent resilience of co-operatives to the Great Recession has strengthened their presence in the world economy, and has aroused the interest in this and other forms of workplace democracy once again (Birchall and Ketilson 2009; Lansbury 2009; Birchall 2013). Some recent philosophical work on economic and workplace democracy has also contributed to the debate (Hsieh, 2008; Schaff, 2012; Perry, 2014; Anderson, forthcoming; Landemore and Ferreras, unpublished; see also the essays in Gosseries and Ponthiere, 2008; and O'Neill and Williamson, 2012: Part II).

Among those who have championed the idea, some have called into question the consistency between the widespread commitment to democracy in the state and the skepticism with which its extension to the workplace is nowadays received. Indeed, it has been argued that firms' decisions influence workers' lives as much as governments' decisions; that managers have as much power over workers as public officials have over citizens; or that large companies influence the society as much as the state does.³ From this point of view, non-democratic firms have sometimes been depicted as autocratic institutions within which the economy is centrally planned, freedom of movement and speech is heavily constrained, and failure to obey can result in instant exile.⁴

1. Democracy is minimally defined throughout the paper as a form of collective decision-making that gives a binding say to all the affected and/or subjected individuals on a roughly equal basis.

2. For some evidence in the US farming context, see United States Department of Agriculture (2004).

3. For references, see section 4, in which these arguments are discussed.

4. Noam Chomsky (1998: 19) has expressed this view as follows: "What kind of freedom is there inside a corporation? They're totalitarian institutions—you take orders from above and maybe give them to people below you. There's about as much freedom as under Stalinism."

However, there may be good reasons against democracy in the workplace that do not apply to the state, and the converse may also be true. Indeed, it has been often claimed that firms and states are too different for the analogy between them to work. For example, it has been argued that firms are voluntary associations while states are not, that firms are for-profit while states are not, and that firms are meritocratic while states are not.⁵

The goal of this paper is to analyze the structure, applications, and limits of the analogy between states and firms that is often used to argue for workplace democracy, i.e. what Joshua Cohen (1989) has labeled as the parallel-case argument for workplace democracy.⁶ According to this argument, firms and states have a number of similarities that make any argument against workplace democracy plausible either in both realms or in neither realm. In this paper I will advance a qualified defense of the parallel-case argument. As I shall argue, firms and states are analogous regarding two salient features (their internal effects and the exercise of power within them). In addition, I will address a number of potential differences between them (regarding the voluntariness of their membership, the narrowness of their goals, and the toughness of the competition they face) that may block the analogy on which the argument is based. I will contend that, while relevant, these differences are often overdrawn, for they are of degree, not of kind. I will conclude that the parallel-case argument provides a sound, yet qualified and non-decisive, basis in favor of workplace democracy.

The paper is divided into five further sections. Section 2 briefly defines workplace democracy and illustrates the parallel-case argument by applying it to two common arguments against workplace democracy. Section 3 unfolds the structure of the argument. Section 4 discusses two salient analogies, while section 5 tackles three potential differences between firms and states. A conclusion closes the paper.

2. THE PARALLEL-CASE ARGUMENT AT WORK

This section has two goals. It firstly introduces the definition of workplace democracy that will be used throughout the paper. It then illustrates the parallel-case argument by applying it to two influential arguments against workplace democracy. (Those who are familiar with these issues may want to directly turn to section 3). Workplace democracy is defined as follows:

5. For references, see section 5, in which these arguments are discussed.

6. The argument has been most recently employed by Schaff (2012), and assessed by Landemore and Ferreras (unpublished). The classic references defending it are Walzer (1983) and Dahl (1985). For a detailed critique, see López-Guerra (2008).

Workplace democracy: A form of managerial organization in which workers have control rights over the management of the firm.

Workplace democracy has developed into many different forms since its nineteenth-century origins, including Robert Owen's cooperative experiments, the Israeli kibbutzim, the German co-determination system, and the US Employee Stock Ownership Plans (Dow, 2003; Hansmann, 2000; Wilkinson *et al.*, 2010: part III). The above definition is, thus, a fairly minimal and inclusive one. Even though it rules out forms of participation that are limited to employee information, communication, and/or consultation, it leaves open a large number of issues regarding the goals, procedures, and boundary conditions of workplace democracy. It also leaves open the relationship between ownership and control rights. Hence, under this definition workers are not required to have a share in the ownership of the firm in order to be granted control rights. They may be granted control rights either *qua* owners or *qua* workers.

The best-known instance of the first alternative is co-operativism, in which workers—and only workers—have equal control rights and supply capital, e.g. via debt contracting or by drawing upon their own savings.⁷ Co-determination, on the other hand, provides the closest instance of the second alternative, in which workers are granted control rights without making any equity investment in the firm.⁸

Many arguments have been advanced in favor of these and other forms of workplace democracy.⁹ Unlike other arguments, the appeal the parallel-case argument is that, by tracing a tight analogy between firms and states, it moves the debate on the desirability of workplace democracy to the political realm, in which the desirability of democracy is taken for granted. In addition, since some of the arguments that are used nowadays against workplace democracy are very similar to arguments that were once used against democracy in the state yet are now seen as unacceptable and anachronistic, the parallel-case argument suggests that we may be using such arguments uncritically.

7. However, as Elster (1989) recalls, only rarely we find cooperatives so-defined, for non-working owners, non-owning workers, and unequal distribution of shares are common.

8. The best-known case of co-determination is the German system (Dow, 2003: chapter 4; Fitzroy and Kraft, 2005). In force since 1976, it makes compulsory for all limited liability firms with over 2,000 workers to have a supervisory board with ample powers (e.g. the approval of the annual budget or the ratification of important investments) in which both shareholders and workers are represented on a "near-parity" basis (because exclusively the shareholders elect the chairman of the board, who has a tie-breaking vote). For a theoretical model of co-determination with a more equal distribution of control rights among shareholders and employees, see Ferreras (2012).

9. For overviews of recent normative debates, see Dow (2003: chapter 2), Hsieh (2008), and González-Ricoy (2010).

In order to illustrate this, let us assume for a moment that the argument is sound, and briefly apply it to two usual arguments against workplace democracy, namely, that employees often lack the expertise required to run a firm (call it the *epistemic argument*) and that only shareholders should have control rights for they are the only suppliers of capital (call it the *argument from capital supply*). (*Just to be clear, in this section I will not assess the merits of the parallel-case argument, something that will only be done in the next section. I only show how the argument could be used if it were sound.*)

2.1. The epistemic argument

Firms are complex institutions that operate in constantly changing economic environments. Their management involves decisions about investment policies, production engineering, contracting, compensation, and budgetary planning, among many other technical and complex issues. Why, then, should workers be granted a say in their governance when they often lack the expertise required to make informed decisions about such issues? As an executive commented, “What? And let the monkeys run the zoo!” (quoted in Christie, 1984: 115). According to the epistemic argument,

- (A1) Complex institutions should not be governed by those who lack the expertise to govern them sufficiently well (i.e. to at least some specific level of competency).
- (A2) Firms are complex institutions and workers lack the expertise to govern them sufficiently well.

Therefore,

- (A3) Firms should not be governed by their workers.

However, consistently extended, (A1) allows for an analogous criticism of democracy in the state. Put simply,

- (A4) States are also complex institutions and not all citizens have the political expertise required to govern them sufficiently well.

Therefore,

- (A5) States should not be governed by all their citizens.

As we shall see below in section 3, it is possible to resist (A5) by claiming that (A1) applies differently to firms and states due to certain relevant differenc-

es between them. For example, efficiency may be crucial in the firm yet not in the state and, accordingly, expertise may be crucial in governing the firm yet not in governing the state. Since the goal of this subsection is just to illustrate how the parallel-case argument could be employed if it were sound, let us assume that (A1) applies equally to both domains. Two possible reactions follow to (A5). On the one hand, it is possible to accept (A1)-(A5). This was common until not so long ago. For instance, in defending the restoration of suffrage restrictions in France right after the Thermidorian coup of 1794, Boissy d'Anglas (1795) famously stated that "we must be governed by the best, and the best are the more educated." Jason Brennan (2011) has recently argued similarly. However, on the other hand, it is possible to deny that it is permissible to disenfranchise some group of voters regardless of how competently they cast their vote, thus rejecting (A5), which most of us would nowadays do. Now, from the latter option it follows that, if (A2) is true, then

- (A6) (A1) should be rejected, i.e. expertise should not be a necessary condition for the governance of complex institutions.

It may be further replied that (A1)-(A5) is too radical an argument, for some degree of insulation of expertise from democratic control may not be at odds with political equality, as the insulation of central banks and constitutional courts from parliamentary decision-making in most democracies proves. This is surely a controversial argument, since it might be argued that the insulation of expertise from democratic control does pose a constraint on political equality, even though such constraint may be justified for reasons other than political equality. However, for present purposes, it is irrelevant whether the argument is sound or not. For, even if it were, it would also apply to firms, given that certain tasks can also be insulated from workers' control in democratic firms. Jeffrey Moriarty (2007: 344) has made the following claim along those lines: "It would be just as unwise to allow employees to elect their firm's Chief Financial Officer as it would be to allow citizens to elect their country's Chairman of the Federal Reserve Board."

The parallel-case argument does not imply that democratic firms ought to include this sort of insulation. It only shows that the scope and limits of democracy and the precise mechanisms of accountability that are to be used are as up for grabs in democratic firms as they are in democratic states. As Walzer (1983: 302) points out, "in a developed economy, as in a developed polity, different decisions are made by different groups of people at different levels of society. The division of power in both cases is only in part a matter of principle; it is also a matter of circumstance and convenience."

2.2. The argument from capital supply

The parallel-case argument can be similarly applied to the argument against workplace democracy from capital supply, according to which workers should not be granted control rights over the governance of the firm because they supply labor but not capital. Workers can always choose to work for democratic firms, take over their own firm in case it goes bankrupt, or try to get a majority of its voting shares. As Nozick (1974: 250) put it, “persons may form their own democratically-run cooperative firms. It is open to any wealthy radical or group of workers to buy an existing factory or establish a new one, and to institute their favorite microindustrial scheme; for example, worker-controlled, democratically-run firms.” In capitalist firms, however, shareholders supply capital. Accordingly, only they should govern the firm. In short, according to the argument from capital supply,

- (B1) Only those who supply capital should govern the firm.
- (B2) Workers supply labor but not capital to the firm.

Therefore,

- (B3) Workers should not govern the firm.

Before turning to the parallel-case argument against (B3), the following caveat is required. Even if we accepted that only shareholders should have control rights over the governance of the firm, it might be argued that workplace democracy need not trump such rights, for it need not be compulsory. True, some (for example, Cohen, 1989) have argued for an inalienable right to workplace democracy. Yet it might be argued that workplace democracy can be implemented gradually and voluntarily, by means of providing legal advice, tax benefits, or direct subsidies to democratic firms, rather than, say, through expropriation or prohibition of non-democratic ones (see Bowles and Gintis, 1996: 66). Now, even when workplace democracy is not compulsory, a rationale is still required to justify why the state should promote democratic firms at the expense of non-democratic ones. To be sure, there is a notable difference between using state coercion to ban non-democratic firms, on the one hand, and using its public resources to promote democratic firms at the expense of non-democratic ones. However, in both cases public means are used to benefit one managerial option at the expense of the other. Hence, the argument from capital supply still poses a potential threat to the justification of a non-mandatory-yet-publicly-promoted workplace democracy.¹⁰

10. I am grateful to Joseph Mazor for pressing me to clarify this.

Let us now go back to premise (B1), according to which supply of capital implies exclusive control rights over the governance of the firm. As in the case of (A1), consistently extended, (B1) leads to a similar criticism in the political realm. The following one:

- (B4) Not all citizens contribute equally to the revenue of the state, if at all.

Therefore,

- (B5) Not all citizens should govern the state.

Again, assume that firms and states are similarly enough for (B1) to apply to both realms. If so, we are again faced with two options. On the one hand, we can accept (B5). This has been a usual way to argue for property and tax qualifications for voting throughout history. John Jay's "favorite maxim", according to which "those who own the country ought to govern it", largely expressed what was common sense until nineteenth—and twentieth—century extensions of the franchise (Jay, 1833: 70). Few would accept (B5) nowadays though. Now, if we reject (B5), then it follows that

- (B6) Premise (B2) should also be rejected, i.e. supply of capital should not be a necessary condition to govern the firm.

3. THE STRUCTURE OF THE PARALLEL-CASE ARGUMENT

As we have just seen, by tracing a close analogy between firms and states, the parallel-case argument pushes the debate on the desirability of democracy *in the workplace* to the political realm, in which democracy is the default normative position. Further, since some of the arguments that are used nowadays against democracy in the workplace, such as the epistemic argument and the argument from *capital supply*, closely resemble arguments that were once used against democracy in the state yet few would accept nowadays, the argument suggests that we may be using such arguments uncritically. In this section I unfold the structure of the argument.

Robert Dahl (1985: 111) provides the best-known account of the parallel-case argument, according to which "If democracy is justified in governing the state, then it must *also* be justified in governing economic enterprises; and to say that it is *not* justified in governing economic enterprises is to imply that it is not justified in governing the state." Of course, the second sentence is redundant. It is just a different yet logically equivalent way to express the material conditional stated in the previous sentence, namely, that

Parallel-case argument: If democracy is justified in governing the state, then it is justified in governing economic enterprises.

Further, even though Dahl formulates it in merely conditional terms, his discussion of the PCA favors a biconditional conclusion. As López-Guerra (2008: 15) points out, it would be certainly awkward if Dahl agreed that democracy could be justified in the workplace yet not in the state, as mere conditionality implies. Even though the previous definition of the argument is enough for the goals of this paper, the following modified version of it follows:

Strong parallel-case argument: Democracy is justified in governing the state if and only if it is justified in governing economic enterprises.

What links the antecedent and the consequent is that economic enterprises and states are taken to be analogous in some morally relevant sense. The parallel-case argument is thus an analogical argument. It refers to some similarities between two objects or systems of objects—namely, firms and states—in support of the conclusion that some further similarity exists (see Bartha, 2010: chapter 1). It unfolds as follows:

- (C1) Economic enterprises are similar to the state regarding certain features.
- (C2) Such features are individually sufficient to justify democracy in governing the state.

Therefore,

- (C3) Such features are individually sufficient to justify democracy in governing economic enterprises.

For this version of the argument to avoid being invalid, at least one further condition needs to obtain. In addition to their similarity regarding certain features that are sufficient to justify democracy in the state, firms and states ought to be similar regarding the absence of a number of aspects that may block the justification of democracy in either realm. For example, it may be the case that being subject to certain form of power by public officials is sufficient to justify democracy in the state, and that managers exercise the same sort of power in the firm. However, it may also be the case that democracy is inappropriate to govern firms because of the stiff competition they face in the market, while it is not to govern the state because states do not face such competition in the international sphere, and that this difference is sufficiently strong to override the similarity regarding the sort of power exercised in both spheres.

In the next section I will consider two potentially relevant similarities between firms and states —regarding their effects and regarding the standing of their members. In the next one, I will turn to three potentially defeating differences between states and economic enterprises. Before turning to these similarities and differences, a caveat is nonetheless in order.¹¹

The plausibility of the conclusions drawn from the argument depends on the moral relevance of the similarities and differences under consideration for the justification of democracy in either realm. Hence, it might be argued that the features considered below in this section —even when similarly present in firms and states— are irrelevant for the justification of workplace democracy because different governance schemes (notably, workplace democracy and political democracy) ought to be assessed according different moral criteria. An argument of this type has been advanced by López-Guerra (2008), who concedes that firms and states might be similar regarding one of the features that will be considered below, namely the exercise of power within them. Yet, he argues, economic justice, and not the exercise of power, should be the criterion employed in assessing the organization of the firm. Accordingly, the parallel-case argument fails because it overlooks the possibility that certain features that are morally relevant for the assessment of some governance schemes may be irrelevant —or not relevant enough to override some further differences that are morally more relevant— when assessing other schemes.

Two replies can be advanced. The first one is that López-Guerra's argument is compatible with the argument presented in this paper. The reason for this is that here I assume a pluralistic view of the values that are morally relevant to the assessment of democracy. Hence, as it will become apparent immediately below, my goal is not to assess if the features that I consider below are morally relevant for the justification of democracy. I make the normative assumption that they are, and that they need to be balanced against each other (something that I do not attempt to do here either). My goal here is rather to analyze if such features are similarly present in firms and states. If the normative assumption turned out to be wrong, then the conclusions drawn from the present analysis would have to be reconsidered.

The second reply is that, even if we accepted López-Guerra's argument, the parallel-case argument could still hold. López-Guerra seems to believe that, if we prove that the exercise of power is not a relevant moral criterion (or not relevant enough to override some other criteria), we then also prove that the argument is invalid. But this is because he explicitly assumes that

11. I am grateful to a referee for this journal for pressing me to introduce this caveat.

the parallel between firms and states has to be based on the exercise of power. To be sure, this has been the main basis in the existing literature, in which it has been assumed that democracy should be equally applied to firms and states because the same sort of power is exercised in both realms. However, there is no reason why the parallel-case argument could not be based on the similarity between firms and states regarding some other moral criteria (for example, how profoundly the decisions made by firms and states affect workers and citizens, respectively). And, once we accept this, it may be the case that firms and states are similar enough regarding these further criteria to make the argument work.

4. SIMILARITIES

Firms and states are similar in a number of ways.¹² However, not *all* the features that firms and states share are equally suitable to be included in this category. These should be limited to those features that may satisfy premise (C2), i.e. those features that may be sufficient for the justification of democracy in the state. Now, different normative theories of democracy will provide different accounts of which precise features count as sufficient in justifying democracy in the state. For instance, pure instrumentalist theories will consider only process-independent features, such as welfare maximization or the protection of fundamental rights, while non-instrumentalist theories will look at process-related features, such as the exercise of power by public officials.

In the remainder of this subsection I will consider two similarities that have dominated recent debates and that are plausible candidates to justify democracy in the state. More specifically, I will briefly consider, first, the external and internal effects of firms and states and, second, the power exercised by managers and public officials.

4.1. Effects

Let us begin with one of the several principles by which democracy has been justified in the state. According to the *principle of all-affected interests*, all which interests are affected by a decision ought to have a say in that decision.¹³ Since the goal of this paper is not to consider which principles may

12. Here I refer to evaluative similarities, i.e. similarities in the values relevant to their assessment, rather than to non-evaluative similarities (e.g. they both are ways of distributing decision-making powers between individuals).

13. For a discussion and defense of the principle, see Goodin (2007). For a criticism, see Saunders (2012).

justify democracy in the state but to assess whether such principles apply equally to firms and states, let us assume that the principle of all-affected interests is sufficient to justify democracy in the state. In considering whether it applies equally to both realms, we need to look at those individuals that are affected by decisions made by firms and by states and the extent to which their effects are similarly pervasive. We can distinguish between two sorts of effects —namely, external effects (i.e. effects on individuals who are members of the two sort of institutions under consideration) and internal ones (i.e. effects on outsiders). Even though external effects turn out to be irrelevant for the issue at hand, let me briefly show why before turning to internal ones.

It has been often argued that firms' decisions have a pervasive influence beyond the limits of the firm, both social and political.¹⁴ Further, it has been claimed that such influence is as pervasive as the influence of states —if not more— in the case of large companies. For instance, in 1999, General Motors' annual revenue was larger than the revenue of the Netherlands, Exxon Mobil's revenue larger than Spain's, DaimlerChrysler's revenue larger than Canada's, and so on (Chowla, 2005: 3). As such, large companies' social and political influence often resembles, if not exceeds, that of states.

However, this analogy is irrelevant for the issue at hand for at least two reasons.¹⁵ First, assuming that the analogy holds, it holds only —or to a much greater extent— for big businesses. The influence of small and medium businesses, by contrast, is not comparable to the influence of states. Second, even assuming that it holds for *all* firms, it does not have a bearing on the justification of democracy, neither within the state nor within firms. Under the principle of all-affected interests, all stakeholders, and not only workers, would have to be granted a say over firms' decisions. Similarly, aliens who are affected by the externalities of the state, and not only citizens, would have to be granted a say over its decisions. Accordingly, stakeholder democracy and global democracy would obtain, rather than democracy *within* the firm and *within* the state.

Consider now the more interesting case of internal effects. It can be argued that decisions made by managers affect workers as much as decisions made by public officials and elected officials affect citizens. On the one hand, firms' decisions can affect workers and their families directly through day-to-day commands or the setting up of the working conditions. For example, in Europe almost as many employees die on average due to fatal accidents

14. As Néron (2010: 336) has put it, “[firms] control vast human, organizational, and financial resources, and labor; they influence national governments and local communities; and they support (directly and indirectly) everything from education to the arts and sports.”

15. I am grateful to two referees for this journal for pressing me to clarify this.

in the workplace as citizens die due to intentional homicide.¹⁶ On the other hand, firms' decisions can affect workers indirectly, as a side effect of strategic decisions such as production planning or relocation. The relevance of these internal effects is enhanced by two further facts. First, workers spend one third of their adult lives in their workplaces, probably more time than anywhere else. Second, given that work is a central source of self-esteem in modern economies, these effects do not have a merely instrumental importance to workers. They are also intrinsically important.¹⁷ In short, internal effects of firms' decisions provide —assuming that the principle of all-affected interests suffices to justify democracy in the state— a robust candidate to ground the parallel-case argument.

4.2. Standing

In defining the similarities between firms and states, most uses of the parallel-case argument have not focused on the influence of firms' decisions in contrast with the influence of states. Rather, they have focused on workers' standing in relation to firms as analogous to citizens' standing in relation to the state and, notably, on the power exercised by managers and public officials.¹⁸ This is a feature that is often seen, at least in the state, as sufficient to justify granting control rights to those over whom such power is exercised (and, again, I will assume that this is the case). Accordingly, it is not very surprising that this type of the argument has been dominant.

It is not very controversial that employees are subject to the power of their employers. In contrast to self-employment, in which workers exchange the product of their labor in the market, rather than their labor force, the very point of the employment relationship is the subordination of the worker to the command of the employer. In large firms, managers rather than owners exercise command as the representatives of the latter in the daily running of the firm. Managers, thus, have power over employees because they have the ability to make the latter perform actions that they would not otherwise perform. What is controversial, then, is not so much whether employers and managers have power over employees. It is clear that they do. And almost as matter of definition, since managers' ability to issue directives to which employees have to conform is a core feature of the employment relationship.

16. In the EU there were 2.5 fatal accidents per 100,000 persons employed in 2008 and 3.5 intentional homicides per 100,000 inhabitants in 2011. See Eurostat (2012: 190) and United Nations Office on Drugs and Crime Homicide statistics, at <http://www.unodc.org/unodc/en/data-and-analysis/homicide.html> (accessed July 6, 2013).

17. On normative issues related to work and self-esteem, see Schwartz (1982), Arneson (1992), and Moriarty (2009).

18. Dahl (1985) is the classic reference relying on the power exercised within the firm.

What is controversial is whether such power is similar to the power exercised by elected representatives and officials in the state. There are three potential differences between firms and states that might call into question that they are.

The first potential difference is that the power exercised by employers is more heavily constrained than the power exercised by elected officials. This could be the case because employers and employees sign a labor contract at the outset of the relationship that clearly specifies the terms under which the relationship will be conducted. By contrast, citizens and elected governments do not sign any such contract. True, in democratic countries citizens elect their representatives. But the latter enjoy ample discretion once they have been elected. They are not subject by binding instructions from the former, or by their own party manifesto. (It might be argued that party manifestos are contracts, but this is at most metaphorical because, unlike labor contracts, they are not legally binding). Accordingly, while employers' exercise of power is heavily constrained (by the employment contract), the exercise of power by elected officials is not.

However, this difference is overdrawn. Neither the discretion of elected officials is completely unconstrained, nor is the discretion of managers completely constrained. In the state, elected officials are legally constrained by vertical and horizontal forms of accountability. First, they are subject to regular elections, in which they need to be reelected. This poses a *de facto* constraint on the extent to which they can deviate from their electoral promises and party manifestos while in office. Second, their power is legally constrained by the constitution and the checks and balances of the other branches of the state. Managers in firms, on the other hand, enjoy ample discretionary powers beyond the terms of their employment contracts because such contracts turn out to be incomplete when they are applied to concrete cases and unforeseeable contingencies. Since it would be impossible or prohibitively costly to anticipate every detail and contingency at the outset of such contract, and since some flexibility is desirable to adequately address such contingencies, employers are unavoidably granted ample discretion to issue commands.¹⁹

A second potential difference is that —unlike citizens in non-democratic states, who do not have a say over decisions imposed upon them by public officials— employees in non-democratic firms do have a say over decisions imposed upon them. This is because they can elect the public officials who regulate the exercise of power in the workplace, and who decide, for that matter, whether the workplace should be democratized or not. Once employees

19. I fully develop this argument in González-Ricoy (2014: 244-248).

get their say in more general elections and are thus able to shape how corporations should operate, it may be argued, the case for democracy in the workplace turns becomes much weaker than the case for democracy in the state.

This, however, does not make the power exercised by non-democratic managers of firms operating in democratic countries irrelevant. To see why, consider the case of a country in which democracy applies at the state level yet not at the municipal level. I take it that the fact that citizens can elect public officials at the state level does not make the power exercised by public officials at the municipal level irrelevant, even if the latter have to exercise their power within the democratic limits imposed by the former, just as it does not make the case for municipal democracy irrelevant. Analogously, the fact that workers have a say at the political level certainly makes a difference for the issue at hand, since it constrains the power that managers can exercise upon them. However, it does not make such power innocuous as far as managers enjoy *some* discretion (something that, as pointed out before, is intrinsic to the employment relationship), and it does not make the case for democracy in the workplace completely irrelevant as a result.

A third potential difference is that the power exercised in the firm is more easily avoidable than the power exercised by public officials. As Arneson (1993: 139) has argued, employees can “generally escape the reach of ... unwanted policies by quitting one’s job and taking another.” Citizens, by contrast, cannot leave their country and enter another one so easily, if at all. The sort of power to which employees and citizens are subjected is, thus, very different. This is an important potential difference that might block the analogy, and it has been extensively discussed in the relevant literature. For now, however, let us put it aside, for it will be discussed in some detail immediately below in section 5.1.

5. DIFFERENCES

In the previous section I have argued that the similarities between firms and states regarding their internal effects, as well as their similarities with regard to the exercise of power within them, are good candidates to ground the parallel-case argument. Assume that this is correct—or that some further similarities between firms and states exist, and that such similarities are, other things being equal, sufficient to favor democracy both in the state and the firm. Even if that is the case, however, these similarities only provide *pro tanto* reasons in favor of workplace democracy. Further differences between the firm and the state may end up overriding them, thus blocking the justification of workplace democracy on balance. In this subsection I will focus on three potential differences. First, firms are voluntary associations while

states are not. Second, firms have well-defined purposes while states are open-ended. Third, firms face stiff competition by other firms while states do not face a similar competition by other states.

5.1. Voluntariness

The potential difference that has dominated the debates about the parallel case is that firms are voluntary associations while states are not. As Arneson (1993: 139) claims, “The most significant disanalogy between states and firms is voluntariness.” The reason for this, according to Arneson and others (Narveson, 1992; see also Dahl, 1985; Mayer 2000; Hsieh, 2008; Cordelli, unpublished), is that workers are entitled to leave economic enterprises at will, while leaving the state may be impossible or very costly. Two problems with how this debate has proceeded are (i) that the notion of voluntariness is rarely made explicit in full and (ii) that it is unclear whether the lack of exit rights necessarily entails that an association is involuntary (which most of the literature about the parallel-case argument assumes). Here I will not attempt to clarify these two problems. Rather, I will assume, following the relevant literature, that exit rights and the ability to exercise them without incurring excessive costs are necessary to deem an association voluntary. When the members of an association lack exit rights, or the costs of exercising them are unbearable (say, because of the absence of acceptable alternatives), then their agreement cannot be deemed fully voluntary.²⁰ In what follows, I will accordingly limit myself to discussing the potential differences between firms and states with this regard.

The basic reason why exit rights are deemed so crucial for voluntariness is that, when the members of an association are entitled to leave it without incurring excessive costs, by remaining inside of it they are taken to consent to the terms of the association. From this standpoint, firms might have been involuntary associations in nineteenth-century England, when Master and Servant Acts were in force and employees were criminally prosecuted for quitting their jobs. And involuntariness may sometimes persist nowadays in monopsonistic labor markets, or in markets of forced labor.²¹ However, in free, competitive, and fully clearing labor markets, so the argument often goes, firms are voluntary associations because employees are entitled to quit

20. Notice, however, that the sort of involuntariness that results from the lack of exit rights is different from the sort of involuntariness that would render a contract nonbinding, as Scanlon (2000: 245) suggests. The mere absence of exit rights does not exempt the parties, thus, from their duty to honor their agreements.

21. According to the ILO's Forced Labour Convention No. 29, forced labor is all work or service that is exacted from any person under the threat of a penalty and for which the person has not offered herself voluntarily.

at will.²² By contrast, states are involuntary because exit is impossible or prohibitively costly.

Two important implications follow. The first one has to do with the analogy between firms and states regarding their internal effects and the exercise of power within them. While citizens can hardly escape such effects as well as the exercise of power by public officials, employees can generally avoid them by terminating their employment contracts. Further, employees can use such freedom as an implicit yet ever-present threat against their employers. As a result, the latter may well *ex ante* modify their behavior so as to incorporate the interests of their employees, thus minimizing the possibility of such freedom being exercised, thus reducing the employee turnover rate. This is not to say that freedom to exit completely rules out employers' power over their employees, or that the latter are not be affected by the decisions of the former any longer. Rather, it means that employees are affected and subjected by such decisions very differently, and to a lesser extent, from how citizens are.

The second implication follows from the first one. As I have argued in the previous section, the parallel-case argument can be grounded on the internal effects and the power exercised within firms and states being similar. Now, if they are not —because, unlike states, firms are voluntary associations, which members are entitled to join and leave at will— then the argument for workplace democracy based on the analogy turns out to be blocked. As Bowles and Gintis (1993: 97) put it, “if the capitalist economy is a sphere of voluntary private interactions, what is there to democratize?” Jan Narveson (1992: 53-54) nicely summarizes these two implications:

“If a firm doesn't like the way you do your job, can it send men with guns who will put you in prison if you don't do it the way the boss says? ... It is fundamental to politics that political association is not essentially voluntary ... Once a gathering is plainly voluntary, then there simply is no case for imposing “democratic” structures and procedures on it.”

However, this contrast between firms and states is overdrawn. Even though it is generally *more costly* to leave one's country than to leave one's job, the difference is one of degree, not of kind, for leaving one's country is, at least formally, as possible as leaving one's job. True, leaving one's country is very costly. It includes serious obstacles such as closed borders, linguistic barri-

22. As a referee for this journal has suggested, markets might not need to fully clear for entry and exit from firms to be voluntary, provided that workers are sufficiently protected from unemployment, e.g. through employment benefits or a basic income. I have considered this alternative argument for the voluntariness of firms in “The Republican Case for Workplace democracy”, section 3. Here, however, I limit the analysis to the stronger argument according to which, even in the absence of such protection, entry and exit from firms in free and competitive markets is voluntary.

ers, and travel expenses, in addition to the fact that moving from one country to another usually implies changing jobs, while one can change jobs without changing countries. However, leaving one's job does not go without sacrifice either.

Briefly consider the following reasons.²³ First, imperfect labor markets do have involuntary unemployment, which makes it costly for employees to quit provided that they would not be able to find another job easily.²⁴ Second, even if labor markets cleared, there are additional exit costs that can lock-in employees, including the following four: (i) investment costs in developing firm-specific human capital; (ii) integration costs in the network of co-workers, customers, etc.; (iii) searching and transition costs from one job to another; (iv) psychological costs in quitting work altogether provided that work is a relevant source of self-esteem in modern societies.

Third, in addition to these costs, freedom to exit, even when costless, may not be sufficient for voluntariness when the alternatives are not acceptable. To see why, consider the following case:²⁵

A is an employee who toils in a humiliating job and wishes to change jobs. For A, leaving her present job is both available and costless. As the labor market fully clears, A has numerous job alternatives available. However, all these alternatives are as humiliating as her present job. Eventually, A decides not to quit and stays at her present job.

Is A's decision *fully* voluntary? I take it that most of us would respond in the negative (even though, as I said above in note 20, not in a sense that would render A's labor contract nonbinding), which shows that freedom to exit does not always suffice for voluntariness. When the range of options available to us is not acceptable, then formal exit rights, which A in the case above holds, do not suffice for voluntariness.

In short, leaving one's job may be less costly than leaving one's country. However, this is a difference of degree, not of kind, for leaving one's job also has important costs. In addition, when acceptable alternatives are absent, freedom to exit, even when costless, does not suffice for voluntariness (even though, as I said above, not in a sense of "voluntary" that would render the contracts of its members nonbinding). Accordingly, any conclusion on the justification of democracy in the workplace drawn from the parallel-case

23. I have developed these reasons in more detail in González-Ricoy (2014: 239-241).

24. Shapiro and Stiglitz (1984) have shown that, under conditions of imperfect information, this is also the case in perfectly competitive markets, which need a sufficiently large unemployment rate to remain competitive.

25. Serena Olsaretti (1998: 71) advances a similar example.

argument would be less compelling than the corresponding justification of democracy in the state. Yet it would not be forceless.

Consider now two objections.²⁶ First, it might be argued that the mentioned costs, as well as the circumstances in which exit rights may be insufficient for involuntariness, vary enormously across employees and firms. For example, employees who possess scarce and valuable skills may bear lower costs if they quit than employees who lack such skills. However, these differences are also present in the state, in which exit costs are also very unevenly distributed across citizens and different states. Hence, some citizens might find it more costly to leave their country than others, and some countries may make it more costly for their citizens to leave than others. Accordingly, the analogy between firms and states regarding these differences holds and if democracy is justified in the state despite these differences, then it follows that it is justified in firms as well.

Second, even if democracy may be favored both in the state and in the workplace due to their similarly profound and unavoidable effects on citizens and workers, respectively, it may be objected that this argument can be blocked by appealing to the idea of freedom of economic association. Workers, it may be argued, have a right to freedom of economic association that empowers them to bind themselves to agree to obey commands of a non-democratic firm, and they have this power even if their set of valuable alternatives is very limited. I raise two points in response. First, as Joshua Cohen (1989: 48) has claimed in response to a similar objection, while it may be valuable to be able to choose the economic activity in which one engages as well as the parties with whom one associates, there may not be any fundamental interest protected by the liberty to sell labor for a wage and to be subjected to undemocratic command in the workplace instead of, say, working as a member of a co-operative. Second, even if we assume that the freedom of economic association does entail a right to work for non-democratic firms, this freedom does not necessarily override workers' right to workplace democracy. Even though these two rights may sometimes conflict with each other, in the sort of economy envisaged by most proponents of workplace democracy—in which workplace democracy (unlike, say, mandatory schooling) is a right that can be alienated—it is not entirely implausible that both rights co-exist.²⁷

26. I am grateful to a referee for this journal and to Andrew Williams for raising these objections.

27. According to Bowles and Gintis (1996: 66), for example, “to argue against mandatory workplace democracy is to critique a straw man and to elide the fundamental issue, which concerns whether policies promoting workplace democracies are justified in the interest of giving workers the opportunity to participate in these forms of governance.”

5.2. Narrowness of purposes

Consider now a further yet related difference, namely, that firms are justified in having well-defined purposes, e.g. to maximize profits, while states ought to be open-ended, as Phillips and Margolis (1999) have argued. The reason for this is that, unlike firms, states are not voluntary associations. Firms are justified in having narrow purposes their employees can always leave if they disagree with such purposes. States, by contrast, have to remain open-ended because their citizens cannot leave them easily if they disagree. Hence, while it is acceptable for a firm to have certain narrow goals (say, produce and sell copies of the Bible), it is unacceptable for the state to do so.

An implication of this difference for the issue at hand is the following. When the goals of an organization are well defined and disagreement about them among its members is not very profound, the need for a collective decision-making procedure, democratic or otherwise, to set the goals that ought to be pursued is also weaker.²⁸ When, by contrast, goals are subject to more profound disagreements, the need for a decision-making procedure to handle such disagreements is stronger. It thus follows that if states are open-ended and have plural goals while firms have narrow purposes, then the case for democratic procedures in the state is stronger than the case for democratic procedures in the firm.

There are good reasons, however, to resist this clear-cut distinction. Firstly, according to Phillips and Margolis, purpose narrowness is allowed in firms and open-endedness is required in the state due to the fact that the former are voluntary organizations while the latter are not. However, as we have seen in the previous section, this difference is overdrawn, for exit from firms is often costly and the decision to stay, thus, not always fully voluntary (at least under the definition of voluntariness used before, which requires that meaningful exit rights are available and that does not necessarily render employment contracts nonbinding). Now, if firms are not fully voluntary associations, then the case for purpose narrowness becomes weaker and the difference between firms and states regarding the narrowness of their goals becomes less clear-cut.

There are further reasons to call into question that purpose narrowness should be allowed in firms and open-endedness should be required in the state. On the one hand, it is certainly the case that some libertarians have

28. As Przeworski (2006: 312) has put it, democracy presupposes, as a necessary condition, that "Interests or values are in conflict. If they were not, if interests were harmonious or values were unanimously shared, anyone's decisions would be accepted by all, so that anyone could be a benevolent dictator." A similar argument can be found in Waldron (1999) and Valentini (2012).

argued not only that making profits is perfectly respectable for economic enterprises, but also that it should be their only goal (typically, Friedman, 1970). However, this position has not gone without challenge, not least by Corporate Social Responsibility approaches and stakeholder theorists. Indeed, it is widely assumed nowadays that firms should have a diversity of goals, social and otherwise, other than maximizing profits. On the other hand, the requirement of open-endedness in the state can also be called into question. Today, it is widely accepted that states have to comply with a number of narrow goals that constrain their sovereignty, including the fulfillment of human rights and the responsibility to protect their population.²⁹

In short, neither firms ought to have narrow purposes, nor the state ought to be completely open-ended. Of course, this does not imply that firms and states ought to have similar goals (just as different firms have different goals). It rather implies that the difference between firms and states regarding the narrowness of their goals is less clear-cut than it is sometimes argued and that, given that firms should also have plural purposes, the argument for ruling out the use of democratic procedures in their governance turns out to be less compelling.

5.3. Toughness and efficiency

Jeffrey Moriarty (2005) has advanced a further difference that may have a bearing on the assessment of the parallel-case argument, namely, the tougher environment that firms face in the market compared to states in the international realm. In free-market economies, firms face stiff competition from other firms that attempt to drive them out from the market. They face continuous and rapid changes due to the appearance of new technologies and products, changes in consumers' preferences, the introduction of new legislation, periodical economic downturns that make them likely to disappear, and so on. Indeed, the US Census Bureau reports that the one-year failure rate for firms started in 2004 is 23.6 percent and the five-year failure rate for firms started in 2000 is 49.3 percent (Headd, 2010). By contrast, states face a less tough environment. They are much more resilient to changing circumstances, and their downfall is rare or at least rarer than in the case of firms.

Two relevant implications follow from this difference. First, Moriarty claims that managers in firms should be granted extensive powers to face stiff competition in the market, as well as the ability to exercise them fast,

29. A classic contemporary defense of human rights as constraints on state sovereignty is Rawls (1999).

that public officials need not have or not to the same extent. In times of economic downturn, he reckons, managers may need to be able to cut employees' pay, give shareholders smaller returns, or renegotiate contracts with suppliers, provided that some minimal constraints (e.g. safety conditions) are respected. Public officials, by contrast, need not have this sort of power, or not to the same degree. The environment they face is less tough, and dissolution of the state less likely to ever happen. Second, even though this is a point that Moriarty does not make, it may also be claimed that stiff competition and the constant threat of downfall make efficiency, in terms of the ratio of output to input, more important in the firm than in the state.

The bearing of these two implications on the parallel-case argument is that both the need for extensive prerogatives and the crucial importance of efficiency may conflict with democratic decision-making, which may be too slow to adapt to changing environments, and may be less efficient than other decision-making arrangements.³⁰ Accordingly, since the need for extensive prerogatives and the importance of efficiency due to stiff competition is greater in the firm than in the state, democratic decision-making may be less suitable in the former than in the latter. These differences, in turn, may block the parallel-case argument for democracy in the workplace or, at least, make any conclusion drawn from other similarities that firms and states may share less compelling.

This is an important argument for, as the figures above suggest, firms certainly face stiffer competition than states. Three replies can be advanced, though. First, governments also face tough circumstances, and the availability of emergency powers and the importance of efficiency may also be crucial in the governance of the state. As Moriarty acknowledges, the difference between firms and states in this regard is one of degree, not of kind.³¹ Second, as it has been argued above, democracy in the workplace is not at odds with delegation of extensive prerogatives to managers, with the only difference that managers in democratic firms are appointed by workers rather than, or along with, shareholders, and accountable to them. Third, it has been much discussed whether democracy in the workplace diminishes or improves efficiency.³² This issue largely exceeds the scope

30. Classic references on the inefficiency of democratic firms are Jensen and Meckling (1979) and Alchian and Demsetz (1972).

31. As Andrew Williams has suggested to me, it may be argued that competition is not only a fact but also a desirable fact in the economic domain, given the benefits of creative destruction. The same, however, may not be true in the political domain, since the social costs of political bankruptcy are so much higher. While this may entail that the difference is ultimately of kind, I leave it open whether the difference holds, for it implies a moral assessment of the benefits of competition that, regarding the economic domain, is highly contested to say the least.

32. Some have argued that the fact that democratic firms are marginal shows that democratic firms are not efficient, for otherwise they would be created voluntarily. See Jensen and

of this paper. However, it may be too quick to assume that efficiency is at odds with democracy, either in the workplace or in the state. There are good theoretical and empirical grounds to believe that the contrary may be the case, at least under certain circumstances (see Bowles and Gintis, 1993; Parks *et al.*, 2004; Levin, 2006; for overviews of the debates, see Dow, 2003; and Hansmann, 2000).

In short, the difference in toughness that firms and states face is one of degree. It should not be overdrawn when assessing the limits to the parallel between firms and states. In addition, the difference might turn out not to have much bearing on the parallel-case argument, first, because democratic firms are consistent with the sort of powers that tough market competition may require and, second, because, they may not be inefficient in their operation.

6. CONCLUSIONS

If the features upon which the analogy between firms and states is based turn out to be sufficient to justify democracy in the state, then the parallel-case argument provides a plausible reason in favor of democracy in the workplace—a reason, however, that needs to be importantly qualified, I conclude, for a number of reasons. First, while the paper has shown that firms and states are similar regarding their internal effects and the power exercised within them, it has not attempted to demonstrate that these features are morally sufficient to justify democracy in the state. Second, the paper has shown that there are a number of morally relevant differences that could override, or at least undermine, the similarities upon which the analogy between firms and states is based—even though it has also shown that these differences are often overdrawn, for they are of degree, not of kind. Third, further differences that have not been considered here may further undermine the analogy between firms and states, thus blocking the parallel-case argument in favor of workplace democracy. In short, the case for democracy in the workplace, when drawn from the analogy between firms and states, and provided that democracy is justified in the state, is plausible. Yet it is also not as strong as the case for democracy in the state. McMahon (1994: 259) is thus right in acknowledging that “the case for democracy in nongovernmental organizations is weaker than the case for democratic government”, at least, when based upon the parallel-case argument.

Meckling (1979). Elster (1989) has replied that they could be marginal due to endogenous preference formation, adverse selection, discrimination, and externalities.

BIBLIOGRAPHY

- Alchian, A., and Demsetz, H., 1972: "Production, Information Costs, and Economic Organization", *American Economic Review* 62: 777-795.
- Anderson, E. Forthcoming: "Equality and Freedom in the Workplace: Recovering Republican Insights", *Social Philosophy and Policy*.
- Arneson, R., 1987: "Meaningful Work and Market Socialism", *Ethics* 97: 517-545.
- 1993: "Democratic Rights at National and Workplace Levels", in *The Idea of Democracy*, ed. D. Copp, J. Hampton, and J. Roemer, Cambridge: Cambridge University Press.
- Bartha, PFA, 2010: *By Parallel Reasoning. The Construction and Evaluation of Analogical Arguments*, New York: Oxford University Press.
- Birchall, J., 2013: *Resilience in a Downturn: the Power of Financial Co-operatives*, Geneva: International Labour Organization.
- Birchall, J., and Hammond Ketilson, L., 2009: *Resilience of the Cooperative Business Model in Times of Crisis*, Geneva: International Labour Organization.
- Bowles, S., and Gintis, H., 1996: "Is Workplace Democracy Redundant in a Liberal Economy?", in *Democracy and Efficiency in the Economic Enterprise*, ed. U. Pagano and R. Rowthorn, eds. (London: Routledge, 1996).
- 1993: "A Political and Economic Case for the Democratic Enterprise", *Economics & Philosophy* 9: 75-100.
- Brennan, J., 2011: "The Right to a Competent Electorate", *The Philosophical Quarterly* 61: 700-724.
- Chomsky, N., 1998: *The Common Good*, Berkeley, CA: Odonian Press.
- Chowla, P., 2005: "Comparing Corporate and Sovereign Power", *Developments* 1: 1-5.
- Cohen, J., 1989: "The Economic Basis of Deliberative Democracy", *Social Philosophy and Policy* 6: 25-50.
- Cordelli, C., unpublished: "On the Scope of Democracy: the Case of Churches".
- Dahl, R. A., 1985: *A Preface to Economic Democracy*, Berkeley: University of California Press.
- D'Anglas, B., 1795: "Discours preliminaire au projet de Constitution pour la République française", *Le Moniteur Universel*, XXV, June 29.
- Dow, G. K., 2003: *Governing the Firm. Workers' Control in Theory and Practice* (Cambridge: Cambridge University Press.
- Elster, J., 1989: "From Here to There; or, If Cooperative Ownership Is so Desirable, Why Are There so Few Cooperatives?", *Social Philosophy and Policy* 6: 93-111.
- Eurostat, 2012: *Europe in Figures. Eurostat Yearbook 2012*, Luxembourg: Publications Office of the European Union.
- Ferreras, I., 2012: *Gouverner le capitalisme? Pour le bicamérisme économique*, Paris: Presses Universitaires de France.
- Friedman, M., 1970: "The Social Responsibility of Business is to Increase Its Profits?", *New York Times Magazine*, Sep 13.
- Fitzroy, F., and Kornelius, K., 2005: "Co-determination, Efficiency and Productivity", *British Journal of Industrial Relations* 43: 233-247.
- Gold, M., 2005: "Worker Directors in the UK and the Limits of Policy Transfer from Europe since the 1970s", *Historical Studies in Industrial Relations* 20: 29-65.
- 2010: "Employee participation in the EU: The Long and Winding Road to Legislation", *Economic & Industrial Democracy* 31: 9-23.

- Goodin, R. E., 2007: "Enfranchising All Affected Interests, and Its Alternatives", *Philosophy and Public Affairs* 35: 40-68.
- González-Ricoy, I., 2010: "Democratizar la empresa: Un análisis desde la filosofía política", *Revista de estudios políticos* 148: 45-69.
- 2014: "The Republican Case for Workplace democracy", *Social Theory and Practice* 40: 232-254.
- Gosseries, A., and Ponthière, G., eds. 2008: "La démocratie d'entreprise", *Revue de philosophie économique* 8.
- Hansmann, H., 2000: *The Ownership of Enterprise* (Cambridge, MA: Harvard University Press).
- Headd, et al., 2010: *Business Dynamics Statistics Brief 5: What Matters More: Business Exit Rates or Business Survival Rates?*, US Census Bureau, Center for Economic Studies, 2010.
- Hsieh, N., 2008: "Survey Article: Justice in production", *Journal of Political Philosophy* 16: 72-100.
- Jay, W., 1833: *The Life of John Jay*, vol. 1, New York: J. & J. Harper.
- Jensen, M. C., and Meckling, W. H., 1979: "Rights and Production Functions: An Application to Labor-managed Firm and Codetermination", *Journal of Business* 52: 469-506.
- Landemore, H., and Ferreras, I., unpublished: "Towards a Justification of the Analogy Firm/State".
- Lansbury, R., 2009: "Workplace democracy and the Global Financial Crisis", *Journal of Industrial Relations* 51: 599-616.
- Levin, H. M., 2006: "Worker Democracy and Worker Productivity", *Social Justice Research* 19: 109-121.
- López-Guerra, C.: "Against the Parallel Case for Workplace democracy", *Revue de philosophie économique* 8: 11-28.
- Marshall, M. G., and Cole, B. R., 2009: *Global Report 2009: Conflict, Governance, and State Fragility*, Vienna: Center for Systemic Peace.
- Mayer, R., 2000: "Is There a Moral Right to Workplace democracy", *Social Theory and Practice* 26: 301-325.
- McMahon, C., 1994: *Authority and Democracy: A General Theory of Government and Management*, Princeton, NJ: Princeton University Press.
- Moriarty, J., 2005: "On the Relevance of Political Philosophy to Business Ethics", *Business Ethics Quarterly* 15: 453-471.
- 2007: "McMahon on Workplace democracy", *Journal of Business Ethics* 71: 339-345.
- 2009: "Rawls, Self-Respect and the Opportunity for Meaningful Work", *Social Theory and Practice* 35: 441-459.
- Narveson, J., 1992: "Democracy and Economic Rights", *Social Philosophy and Policy* 9: 29-61.
- Néron, P., 2010: "Business and the Polis: What Does it Mean to See Corporations as Political Actors?", *Journal of Business Ethics* 94: 333-352.
- Nozick, R., 1974: *Anarchy, State and Utopia*, New York: Basic Books.
- Olsaretti, S., 1998: "Freedom, Force and Choice: Against the Rights-Based Definition of Voluntariness", *Journal of Political Philosophy* 6: 53-78
- O'Neill, M., 2008: "Three Rawlsian Routes towards Economic Democracy", *Revue de philosophie économique* 8: 29-55.

- O'Neill, M., and Williamson, T., eds., 2012: *Property-owning Democracy: Rawls and Beyond*, Oxford: Blackwell.
- Parks, R.; Kruse, D., and Sesil, J., 2004: "Does Employee ownership Enhance Firm Survival?", in *Employee Participation, Firm Performance and Survival*, ed. V. Perotin and A. Robinson, Amsterdam: JAI Press.
- Perry, F., 2014: "Reducing Racial Prejudice through Workplace Democracy", *Journal of Social Philosophy* 45: 203-227.
- Phillips, R. A., and Margolis, J. D., 1999: "Toward an Ethics of Organizations", *Business Ethics Quarterly* 9: 619-638.
- Przeworski, A.: "Self-Enforcing Democracy", in *Oxford Handbook of Political Economy*, ed. D. Wittman and B. Weingast, New York: Oxford University Press.
- Rawls, J., 1999: *The Law of Peoples*, Cambridge, Mass.: Harvard University Press.
- Saunders, B., 2012: "Defining the Demos", *Politics, Philosophy & Economics* 11: 280-301.
- Scanlon, T., 2000: *What We Owe to Each Other*, Cambridge, MA: Harvard University Press.
- Schaff, K. P., 2012: "Democratic Rights in the Workplace", *Inquiry* 55: 386-404.
- Schwartz, A., 1982: "Meaningful Work", *Ethics* 92: 634-646.
- Shapiro, C., and Stiglitz, J., 1984: "Unemployment as a Worker Discipline Device", *American Economic Review* 74: 433-444.
- United States Department of Agriculture, 2004: *Farm Marketing, Supply and Service Cooperative Historical Statistics*, Washington, DC: USDA, Rural Business, Cooperative Service.
- Valentini, L., 2013: "Justice, Disagreement and Democracy", *British Journal of Political Science* 43: 177-199.
- Waldron, J., 1999: *Law and Disagreement*, Oxford: Oxford University Press.
- Walzer, M., 1983: *Spheres of Justice*, New York: Basic books.
- Wilkinson *et al.*, eds., 2010: *The Oxford Handbook of Participation in Organizations*, Oxford: Oxford University Press.

Rethinking the Good - A Small Taste

LARRY TEMKIN

Rutgers University

Abstract

This article aims to convey a few of the key claims and arguments of my book, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. The article gives an example of a *Spectrum Argument*, and illustrates that such arguments put pressure on the *Axiom of Transitivity*, which holds that for any three possible outcomes or alternatives, A, B, and C, if, all things considered, A is better than B, and B is better than C, then A is better than C. The article distinguishes between two different approaches to understanding the goodness of outcomes, the *Internal Aspects View* and the *Essentially Comparative View*. It suggests that two deeply plausible, but seemingly incompatible, positions underlying the Spectrum Argument, an *Additive-Aggregationist Position*, and an *Anti-Additive-Aggregationist Position*, reflect the Essentially Comparative View, and that on such a view they are not incompatible. The article introduces several widely-held views about neutrality and dominance principles, and shows that some of these views are incompatible. The article contends that various ideals or views that people care about are most plausibly understood as essentially comparative, and notes that one such view, a *Narrow Person-Affecting View*, will be especially difficult to reject in at least some cases. It also illustrates how such a view, like other essentially comparative views, threatens the Axiom of Transitivity. The article concludes by contending that we must seriously rethink our understanding of the good, moral ideals, and the nature of practical reasoning, while recognizing that the way forward is murky, at best.

Keywords: Transitivity, Practical Reasoning, Internal Aspects View, Essentially Comparative View, Narrow Person Affecting View, Spectrum Argument, Additive Aggregation, Good, Better than, Ideals.

This article is based on my Fall 2012 LEAP Lecture given at Pompeu Fabra University. The Lecture kicked off a symposium on my book, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning* (Temkin, 2012), with responses to the book offered by Oscar Horta and Ingmar Persson, fol-

lowed by comments from me on those responses¹. The aim of the Lecture was not to give an overview of the book, which would have been impossible in the time allotted, but rather, as I told the audience, to give a very crude and brief tour of a few of the book's arguments, just enough to give a sense for the sorts of issues the book explores. Correspondingly, this article, like the Lecture from which it is derived, is woefully incomplete and superficial. But, hopefully, some readers will find it sufficiently important and intriguing to turn to the book itself, where a more careful and sustained treatment can be found of the issues broached here, as well as many other issues central to our understanding of the good, moral ideals, and the nature of practical reasoning.

This article is divided into six sections. In section I, I provide a brief introductory remark, and offer a simple example of a *Spectrum Argument*. The Spectrum Argument puts pressure on a widely accepted principle of practical reasoning which may be called the *Axiom of Transitivity*. According to the Axiom of Transitivity, for any three alternatives, A, B, and C, if, all things considered, A is better than B, and B is better than C, then, all things considered, A is better than C.² In section II, I offer some background to some of the issues I discuss, and make some terminological distinctions. In section III, I introduce a distinction between two different approaches to understanding the goodness of outcomes, which I call the *Internal Aspects View* and the *Essentially Comparative View*. I note how two seemingly incompatible positions underlying the Spectrum Argument, which I call an *Additive-Aggregationist Position*, and an *Anti-Additive-Aggregationist Position*, can be seen as reflecting the Essentially Comparative View, and that on such a view they are not incompatible. I also note various considerations against rejecting the Anti-Additive-Aggregationist Position. In section IV, I introduce several widely-held views about neutrality and certain widely-held dominance principles. I show that some of these views are incompatible. In section V, I suggest that various ideals or views that people care about are most plausibly understood as essentially comparative. I focus on a particularly plausible

1. I want to thank Paula Casal and José Luis Martí for inviting me to deliver the LEAP Lecture, for organizing the symposium, and for arranging for the publication of the symposium's papers. I would also like to acknowledge my gratitude to Horta and Persson for their careful and thoughtful attention to my work.

2. Here, I am using "the Axiom of Transitivity" as shorthand for "the Axiom of Transitivity of the 'all-things-considered better than' relation". Elsewhere, I often put my discussions in terms of "the Axioms of Transitivity", where these include the "all-things-considered equally as good as" and "all-things-considered at least as good as" relations as well as the "all-things-considered better than" relation. At times, I may shorten my descriptions and just talk in terms of the "better than", "equally as good as", or "at least as good as" relations. But, unless noted otherwise, if I consider whether one outcome is better, equally as good as, or at least as good as, another, I am considering whether the one outcome is better than, equally as good as, or at least as good as the other *all things considered*.

version of a *Narrow Person-Affecting View*, and note how this view, like other essentially comparative views, threatens the Axiom of Transitivity. In section VI, I conclude with some final remarks.

1. INTRODUCTION AND A SPECTRUM ARGUMENT

In this article, I will be discussing a number of views that are widely taken to be *obviously* true. At first blush this may seem rather odd. Why labor the obvious? The answer, in a nutshell, is that a number of the seemingly obvious views aren't even *true*, much less obviously so! This follows from the simple fact that a number of the so-called "obvious" truths are incompatible with each other. Or so I shall argue anyway. Indeed, on reflection, it turns out that an awful lot of hard work needs to be done to sort out what we really *should* believe in the domains I shall be canvassing. I can't do the required work here, in this article, but perhaps I can say enough to motivate the importance of taking up the task. I tried, in *Rethinking the Good*, to do much of the work in question. The result of that work, I believe, is that we need to significantly revise our current understanding of the good, moral ideals, and the nature of practical reasoning, and that such revisions will have profound practical and theoretical implications. The aim of this article is to provide a small taste of the questions addressed in my book, and what is at stake as we try to answer them.

Let me begin by presenting two very simple questions, and the answers these questions typically provoke.

My first question goes like this. Suppose that you or a loved one are going to have to experience a certain intensity of pain, for a certain duration, or a *little* bit less intense pain for *twice*, or *three*, or *five* times as long. Which alternative do you think would be *better* for you or your loved one?

When I asked that question during my LEAP Lecture, there was *total* agreement amongst the audience of roughly forty people, that the first alternative would be better; that is, that an outcome involving a slightly *more* intense pain would be *better* than an outcome involving a slightly *less* intense pain, *if* the duration of the pain in the outcome with the less intense pain would be two, or three, or five times as long as the duration of the pain in the outcome with the more intense pain.

The audience's responses were very typical. Among audiences around the world, involving 1000s of people over many years, virtually *everyone* thinks the *better* outcome would be the one with a *slightly* more intense pain that lasted *significantly* less long. Indeed, I estimate that over 95% of the people of whom I have asked my question have responded the same way; and, as I usually like to put it, only half in jest, if several people in an audience of a

hundred have answered differently, typically one or two are just being difficult, or figuring it is a trick question, and the other one or two haven't fully understood the question!

My second question goes like this. Suppose that you, or a loved one, are going to live for a long time. Perhaps a *very* long time. And there are two ways your life might go. In one, you will have, on average, fifteen mosquito bites a month for the duration of your life and, in addition, at some point in your life you will have two years of the most *excruciating* torture imaginable—including such things as hot wax under your eyelids, bamboo shoots under your fingernails, electrical shocks to your genitals, and so on. You would be awake 18-20 hours per day, and during *every* waking moment your life would be *much* worse than nothing and you would wish you were dead. However, after the two years of torture, you would be given a pill so that you didn't remember any of the pain. Further, let us suppose that the torture would have no permanent impact on your body or brain, and that there would be no *other* effects of any kind during the remainder of your life, once the two years of excruciating pain was over. In the second way your life might go, there would be no torture of any kind. However, instead of *fifteen* mosquito bites per month for the duration of your life, you would have *sixteen* mosquito bites per month. Bearing in mind that your life might be *very* long, which life would be *better* for you or your loved one; the life with fifteen mosquito bites throughout and two years of excruciating torture, or the life with sixteen mosquito bites throughout?

To this question, all but one member of the LEAP Lecture audience gave the same answer. And I think it is fair to say that many audience members were dumbstruck when someone voted for the position that the life involving two years of torture would be *better* than the life involving one extra mosquito bite a month, if only the two lives lasted long enough!

As before, the reactions of the LEAP audience were very typical. Of the thousands of people to whom I have posed such a question over the years, the *vast* majority of them—again, well over 95% I would estimate—have given the same answer to this question. They think that the life involving *one* extra mosquito bite per month would be better, indeed *much* better, than the life involving two years of excruciating torture, and they think this no matter *how* long the two lives might persist.

As indicated, these two results are *very* robust. But together, they are inconsistent *if* one accepts the Axiom of Transitivity: that if, *all things considered*, A is better than B, and B is better than C, then *all things considered*, A is better than C. To see this, notice that when I asked my first question, I didn't actually say *how* intense the two pains were, nor how *long* they lasted. And I didn't need to! This is because it seems to be a *general* truth that no matter

how intense a given pain might be, and how *long* it lasted, it would be better to have *that* pain than one that was only *slightly* less intense but which lasted *much* longer.

Accordingly, one can imagine a spectrum of lives, each of which would be *very* long and each of which would have, as a persistent background condition, fifteen mosquito bites per month. The first life in the spectrum would also involve *extraordinary* pain (the equivalent, let us suppose, of excruciating torture) lasting for two years, and each subsequent life in the spectrum would involve *slightly* less intense pain than that involved in the preceding life in the spectrum, but the pain would last two, or three, or five times as long as the duration of pain in the preceding life of the spectrum. Moving from the first member of the spectrum to the last, the pain gets slightly less intense though much longer, until *eventually* the pain has decreased *so* much that its intensity is the equivalent of but *one* extra mosquito bite per month, though instead of only lasting *two* years, as the pain did in the first member of the spectrum, the once a month mosquito-like pain extends throughout much, if not all, of the very long life.

The point, of course, is that in accordance with the answer to the first question I asked, most people would agree that, all things considered, the first member of the spectrum would be *better* than the second, the second would be *better* than the third, the third would be *better* than the fourth, and so on. For each pairwise comparison, the life involving fifteen mosquito bites per month and a slightly more intense pain lasting a certain duration would be better, all things considered, than the life involving fifteen mosquito bites per month and a slightly less intense pain lasting two, or three, or five times as long. According to the Axiom of Transitivity, it follows that the *first* member of the spectrum must be *better* than the *last*. But the first member of the spectrum involves a life involving 15 mosquito bites per month and two years of excruciating pain the equivalent of torture, and the last member of the spectrum just involves 15 mosquito bites per month and *many* years of a minor pain that is the equivalent in intensity to *one* extra mosquito bite per month! Thus, as we have seen, most people would *reject* the claim that the first member of the spectrum would be better than the last. Indeed, I have found that most people —though admittedly not *all*— regard such a view as preposterous, if not downright absurd.

It follows that *if* people want to maintain the answers typically given to my two questions above —answers to which, I believe, most people are deeply committed— then they must reject the Axiom of Transitivity.³

3. The first Spectrum Argument challenging the Axiom of Transitivity was developed by Stuart Rachels (1993). Rachels' thinking about intransitivity was sparked by my original article on the topic, "Intransitivity and the Mere Addition Paradox" (Temkin, 1987), but his argument

This is a very striking result. Because the Axiom of Transitivity is one of the key premises underlying *Expected Utility Theory*, and Expected Utility Theory is arguably the central theory underlying game theory, decision theory, and much of modern economics. So, rejecting the Axiom of Transitivity would entail rejecting, or substantially revising our understanding of, game theory, decision theory, and much of modern economics. Since, in many ways, those theories are intended to model our best understanding of practical rationality, rejecting the Axiom of Transitivity would require us to drastically revise our understanding of what it is to be practically rational.

Put differently, the Axiom of Transitivity lies very close to the core of our current understanding of practical rationality. We believe that just as it is irrational to believe both A and not A, or to prefer A to B or believe that A is better than B, all things considered, while *at the same time also* preferring B to A, or believing that B is better than A, all things considered, so, too, we believe that it is irrational to prefer both A to B, and B to C, or to believe both that A is better than B and that B is better than C, all things considered, *while at the same time also* preferring C to A, or believing that C is better than A, all things considered.

As economists would often put it, someone with intransitive preferences is *irrational* and they ought to get their preferences in order! In this context, the “ought” is the strong normative “ought” of individual rationality, implying that rationality *requires* that their preferences be transitive.

It is worth adding that the Axiom of Transitivity is not merely an important theoretical assumption underlying our understanding of ideal rationality and some important academic fields, it plays an integral role in countless cases of everyday practical reasoning, typically without our even being aware of the role it is playing. For example, often when we are faced with a decision between various alternatives with a number of competing factors relevant to our decision, and a significant degree of indeterminacy involved regarding how much to weight each factor, we simplify our decision procedure by focusing on just two alternatives at a time.

against intransitivity was entirely original and at the time it was the strongest argument yet posed against the Axiom of Transitivity. Although I have developed and defended Spectrum Arguments in my own way over many years now, the basic structure of my arguments remains heavily indebted to Rachels's original argument. Rachels' published contributions in this area include Rachels, 1998, 2001 and 2004. Many people have worried about the implausibly long length of life that might be involved in the kind of Spectrum Argument presented in the text. I address such worries in *Rethinking the Good*, but also show that similar arguments can arise involving many different people all living at the same time, rather than a single person living at many times (see chapters 2, 5, and 9 for extended discussion and defense of Spectrum Arguments).

For instance, suppose we have decided to buy a new car, and based on our research we have narrowed our choice down to seven models. At that point, we might test drive the first model, and then test drive the second, and then, taking account of each of the factors that are important to us and how much we care about them—cost, gas mileage, reliability, resale value, ease of repairs, handling, storage capacity, power, handling, comfort, looks, extra features, and so on—we might determine that, *all things considered*, the first model, A, is better than the second, B. In that case, we remove B from further consideration, test drive C, and then decide whether A is better than C. If C is better we remove A, from further consideration, test drive D, and proceed as before.

In this way, we might straightforwardly determine which of the seven models to buy on the basis of a sequence of six direct pairwise comparisons, with the “winner” of each pairwise comparison advancing to a subsequent comparison, and the “loser” being discarded from further consideration. As long as we are confident in each of our pairwise judgments, we will be confident that we have determined the best car for our purposes given our preferences. Moreover, given the many different factors we have to pay attention to, focusing clearly and carefully on the various models just two at a time, we will often be much more confident in any comparative judgments we might arrive at as to which of two cars is better, *all things considered*, than we would be in any absolute judgments about exactly *how* good each of the seven cars were, *all things considered*.

As indicated, this simplifying decision procedure of focusing on just two alternatives at a time is a staple of many practical decisions involving multiple options. But, importantly, this decision procedure *depends* on the Axiom of Transitivity for its legitimacy. After all, we can only confidently remove B from further consideration after determining that A is better than B, *all things considered*, *if* we can be certain that it *couldn't* be the case that there is some third alternative, C, which is both *worse* than B, and yet *better than* A, *all things considered*. For if it *could* be the case that, *all things considered*, A is better than B, which is better than C, which is better than A, then there would be no more reason to remove B from further consideration just because it is worse than A, than there would be to remove A from further consideration given that it is worse than C, or C from further consideration given that it is worse than B. It is the Axiom of Transitivity which presumably “guarantees” that this unfortunate predicament couldn't arise. Thus, as indicated, the Axiom of Transitivity is presupposed, often implicitly and unwittingly, in numerous cases of everyday practical reasoning. Clearly, such reasoning is deeply flawed if the Axiom of Transitivity fails to hold.

I suggest, then, that there is a great deal at stake, both theoretically and practically, if the Axiom of Transitivity fails. And for many years, I argued that

Spectrum Arguments, such as the one given above, as well as various other arguments I developed, gave us good reason to conclude that the Axiom of Transitivity *does* fail. That is, I *used* to claim that we *should* conclude that all things considered better than is *not* a transitive relation. But my earlier claims were too strong, and hence misleading.

What I now think is that over the years I have developed a series of *impossibility arguments*. The Axiom of Transitivity is *one* of the key premises in my impossibility arguments, but it is not the only one. Accordingly, *each* of the key premises of my impossibilities arguments are in play and, if the reactions to the work in this area over the years are any indication, the question of which of the premises should be given up is a difficult one about which people are deeply divided, and about which there is unlikely to be a consensus for years to come.

A second key premise that is in play in Spectrum Arguments is a position I call the *First Standard View: Trade-offs between Quality and Number are Sometimes Desirable*. On this view, in general, it is better to experience more intense suffering for a shorter period of time than less intense suffering for a longer period of time, *if* the difference in the intensity of the two pains is *sufficiently* small, and the difference in their durations is *sufficiently* large.

A third key premise that is in play in Spectrum Arguments is a position I call the *Second Standard View: Trade-offs between Quality and Number are Sometimes Undesirable Even When Vast Numbers are at Stake*. On this view, in general, it would be worse to receive a more intense pain of a significant duration than a much less intense pain of virtually any duration, *if* the difference in intensity of pains is such that the more intense pain of significant duration would have a significant negative impact on one's life, while the less intense pain of longer duration would have little negative impact on one's life.

Each of the Axiom of Transitivity and the First and Second Standard Views is powerfully appealing, and I believe that giving *any* of them up would have deeply implausible implications. So my current position is like that of a juggler, who is juggling a number of very valuable and fragile balls, and he can't hang on to all of them. He has to let at least one of them drop, but can't decide which one. Initially, he may decide to let the first ball drop, and preserve the others. But as the first ball heads towards the ground he thinks he can't possibly let *that* ball drop, so he quickly reaches out to preserve that ball and lets the second ball go, instead. But he then realizes that he can't let *that* ball drop either, so he seeks to save that one, as well, steeling himself to let the third ball drop. But as the third ball gets closer and closer to the ground he realizes he can't bear the thought of losing that ball either, so reaches out to save it with the thought that he'll let the fourth ball go. This process contin-

ues, till he once again finds himself letting the first ball drop. The problem, of course, is that the cost of letting *any* of the valuable balls go seems unacceptably high, so he frantically wants to keep *each* of them in the air, but realizes that that option is ultimately unsustainable.

To a large extent, my book is about determining what various positions stand or fall together, and illuminating both the benefits and costs associated with retaining or abandoning each of the offending premises in my impossibility arguments.

2. SOME BACKGROUND AND TERMINOLOGY

Many believe that giving up the Axiom of Transitivity is not an option. They believe that it is an analytic truth—literally true in virtue of the meanings of the words—that “all-things-considered better than” is a transitive relation. This is the view of John Broome (1991 and 2004), and at one time it was the view of Tom Nagel, Tim Scanlon, and Derek Parfit.⁴ I suspect that this, or something very close to it, is also the view of many economists, for whom the transitivity of the “all-things-considered better than” relation is an unquestioned, and perhaps even self-evident, axiom which *needs* no argument. I think this view is mistaken or, more charitably, deeply misleading.

Since people can use words as they see fit, let me first simply grant that there may be *a* use of the words “all-things-considered better than” such that it must be a transitive relation, by definition. So, if Broome or others want to insist that as *they* use the notion of “all-things-considered better than” the Axiom of Transitivity *is* analytic, there is no point in denying or trying to refute their claim. But then, let me hasten to add that, as Wittgenstein might have put it, meaning is use, and there is another, widely accepted and more normatively significant, usage of “all-things-considered better than”, what I call the *reason-involving* sense of “all-things-considered better than”, according to which to say that A is better than B, all things considered, is to say that from an impartial perspective there is most reason to rank A as more desirable than B taking full account of *all* of the factors that are relevant and significant for making *that* comparison.⁵ And, as I shall suggest next, on *that* notion of “all-things-considered better than”—the reason-involving one—even if it *is* true that “all-things-considered better than” is a transitive relation, it is not an *analytic* truth, rather, it is a truth that turns on substantive facts about the nature and structure of the good.

4. Nagel’s, Scanlon’s, and Parfit’s early views on this topic were conveyed to me during discussions when I was a graduate student (for more on this see my Preface in Temkin, 2012).

5. Ludwig Wittgenstein’s famous contention that “meaning is use” is defended in Wittgenstein 1958.

3. THE INTERNAL ASPECTS VIEW VERSUS THE ESSENTIALLY COMPARATIVE VIEW

To see how the transitivity of the “all-things-considered better than” relation in the reason-involving sense turns on substantive facts about the nature and structure of the good, it will help to consider two alternative models for thinking about ideals in general, and moral ideals in particular, which I call the *Internal Aspects View* and the *Essentially Comparative View*.

Here is one natural and plausible way of understanding the Internal Aspects View. On this view, how good or bad any given outcome is with respect to any given ideal depends solely on the internal features of that outcome. Likewise, how good or bad any given outcome is all things considered will depend solely on how good or bad it is with respect to each ideal. Now this will be a function of how much the different ideals matter relative to each other, and it may, in fact, be a very complex function reflecting various holistic interaction effects between different ideals, but the *key* point is that on the version of the Internal Aspects View that I am now elucidating, ultimately there is a *fact* of the matter about how good or bad each outcome is, and that fact depends solely on the *internal* features of that outcome and the internal relations between them.

So, on the Internal Aspects View, if one wants to assess how good or bad an outcome is, all things considered, it will always be sufficient to consider that outcome *directly*, by *itself*, in terms of all of the factors or ideals that are relevant and significant for assessing the internal features of outcomes. Thus, for example, if one believes that equality is relevant to the goodness of outcomes, one will consider the extent to which equality or inequality is a feature of that outcome, and similarly for other relevant ideals such as justice, freedom, utility, perfection, and so on. One will then give each outcome its due weight, taking account, as necessary, of any relevant interaction effects, in order to arrive at an all things considered judgment regarding the outcome's overall goodness.

The Internal Aspects View allows room for epistemological ignorance about how good or bad any given outcome is, as well as room for believing that facts about the goodness of outcomes may be indeterminate or imprecise, but it is natural to assume that each outcome will have a precise or imprecise degree of goodness or badness that can, in principle, be accurately represented by a number or range of numbers on the real number line. So, for example, in principle it might be a fact that, all things considered, any given outcome might have 1013 “units” or “degrees” of goodness or, alternatively, perhaps there may be no fact as to *precisely* how good the outcome is, but it might still be true that it has between 1003 and 1023 “units” or “degrees” of goodness. For simplicity, in what follows I shall ignore the compli-

cation introduced by imprecision, and assume that each outcome can be given a precise number representing its degree of goodness. But the points I am making could also have been made in terms of ranges of numbers for those who believe that the degree or extent to which an outcome is good or bad is (often) imprecise, and best captured by a range of numbers rather than a single number.⁶

The Internal Aspects View is a natural and plausible way of thinking about ideals and their relation to the goodness of outcomes. It also supports various views that have been thought central to practical reasoning or the assessment of outcomes. For example, it clearly supports the Axiom of Transitivity, since if the number representing A's degree of goodness based solely on A's internal features is higher than the number representing B's degree of goodness based solely on B's internal features—which will be the case if A is better than B—and the number representing B's degree of goodness based solely on B's internal features is higher than the number representing C's degree of goodness based solely on C's internal features—which will be the case if B is better than C—then the number representing A's degree of goodness based solely on A's internal features *will* be higher than the number representing C's degree of goodness based solely on C's internal features—since “being a higher number than” *is* a transitive relation—and hence A *will* be better than C precisely as the Axiom of Transitivity requires.

The Internal Aspects View also supports another principle which many economists and others have regarded as a central principle of practical reasoning, which is often called the Independence of Irrelevant Alternatives Principle (*IIAP*). On *IIAP*, to know how A compares with B it is sufficient to compare them *directly*, as how A or B compares with respect to some third alternative, C, or some other set of alternatives C through N, is *irrelevant* to how A compares with B. As we have seen, on the Internal Aspects View, any outcome A will get a score representing its degree of goodness and that score will be based solely on A's internal features. And similarly for any outcome B. A will be better than, equal to, or worse than B, if and only if its score is higher than, equal to, or lower than B's, respectively. Accordingly, how A compares to B in terms of goodness follows directly from how good each of

6. Some people reject the numerical model entirely. For example, in discussion, both Derek Parfit and Ingmar Persson have conveyed their rejection of any sort of numerical model for understanding the good. But while there are problems with any numerical model, I think this way of thinking about the Internal Aspects View is natural, plausible, and sufficient for my present purposes. I might add that a well-worked-out alternative to such a model has not yet been given. Moreover, I am skeptical as to whether a coherent non-numerical model can be developed which will capture the most important and attractive features of an Internal Aspects View. I briefly touch on this issue at the end of my response to Persson's article (see Temkin, 2014: ??? pp. 18-19 of my final draft).

them is, considered just by itself, and doesn't depend *at all* on how either or both of them compares to some third alternative or some other set of alternatives. Thus, as indicated, the Internal Aspects View supports, or indeed implies, the Independence of Irrelevant Alternatives Principle.

Let me mention just one other principle of practical reasoning which has great plausibility and which is supported by the Internal Aspects View. It is plausible to believe that if two alternatives, A and B, are equally good, then however A compares to some third alternative C, that is exactly how B will compare to C. I call this principle the *Principle of Like Comparability for Equivalents*. It is easy to see how the Principle of Like Comparability for Equivalents holds if the Internal Aspects View is correct. On the Internal Aspects View, for any three outcomes, A, B, and C, how good A, B, and C are will depend solely on their internal features, and each of them will receive a score representing its degree of goodness. If A and B are equally good they will receive the *same* score, so clearly however A's score compares to C's score, that is how B's score compares to C's score.

In sum, the Internal Aspects View has great intuitive plausibility, and it would support and explain a number of other widely accepted views about practical rationality that many have found compelling, including the Axiom of Transitivity, the Independence of Irrelevant Alternatives Principle, and the Principle of Like Comparability for Equivalents. The problem is that despite its great appeal, the Internal Aspects View doesn't reflect the thinking that many people often engage in when assessing outcomes! In particular, as I argued in chapter 12 of *Rethinking the Good*, many of the ideals people value most reflect an *Essentially Comparative View* of moral ideals. This includes especially plausible versions of Utility, Maximin, the Pareto Principle, and the Narrow Person-Affecting View.⁷ On such views, there is no fact of the matter as to how good or bad an outcome is considered just by itself with respect to the ideal in question, or if there is, that fact has no special significance in comparing outcomes with respect to that ideal. Rather, our assessment of how good an outcome is with respect to the ideal in question will depend on the alternative or alternatives with which it is compared. More specifically, on an Essentially Comparative View of ideals, the relevance and significance of the factors for assessing how good an outcome is regarding a

7. Roughly, Utility assesses the goodness of outcomes in terms of how much utility, or well-being, the sentient beings in those outcomes have, Maximin assesses the goodness of outcomes in terms of how well off the worst-off individuals fare in those outcomes, and the Pareto Principle claims that in outcomes involving the same people, one outcome will be better than another if it is better for at least one person and at least as good for everyone else. I'll discuss the Narrow Person-Affecting View more later. As stated in the text, in Temkin, 2012: ch. 12, I argue that in many contexts, the most plausible versions of the ideals in question are Essentially Comparative.

particular ideal may *differ* depending on the outcome's alternative(s), so, in essence, a given outcome may have one value regarding an essentially comparative ideal given one alternative, but a different value regarding that very same ideal given another alternative.

It follows that if an Essentially Comparative View of moral ideals is correct—so, for example, in comparing certain outcomes it is legitimate, as many believe, to assess them in terms of essentially comparative versions of Utility, Maximin, the Pareto Principle, or a Narrow Person-Affecting View—then there is no reason to expect the “all-things-considered better than” relation to be transitive. This is because if the relevance and significance of the factors for assessing an outcome can vary depending on the alternative with which it is compared, then it could well be the case that for any three alternatives A, B, and C, A might be better than B in terms of *all* of the factors that are relevant and significant for making *that* comparison, and B might be better than C in terms of *all* of the factors that are relevant and significant for making *that* comparison, and yet A might *not* be better than C in terms of *all* of the factors that are relevant and significant for making *that* comparison. After all, it could then well be the case that the factors that are relevant or significant for comparing A with C, and which might rightly support the judgment that A is *not* better than C, may *differ* from the factors that are relevant and significant for comparing A with B, or B with C, allowing for the real possibility that *those* factors might rightly support the judgment that A is better than B, and B is better than C.

So, in reflecting on whether or not the Axiom of Transitivity holds, a *key* question is whether the nature and structure of ideals reflects an Internal Aspects View of the sort sketched above, or an Essentially Comparative View of the sort sketched above. And I submit that the answer to this question is a *substantive* matter determined by the nature of the normative domain, it is *not* a terminological matter determined by the *meanings of the words* “all-things-considered better than”! The words “all-things-considered better than” can't dictate the nature and structure of the normative realm. *If* ideals have the structure embodied by the Internal Aspects View as I have characterized it, then, indeed, the Axiom of Transitivity will hold. But *if* at least some ideals have the structure reflected by the Essentially Comparative View—as *might* be the case—then it will not.

I submit, then, that in the face of seemingly compelling arguments that put pressure on the Axiom of Transitivity, we must do the hard philosophical work of facing those arguments head on and determining which, if any, of their premises should be rejected. We cannot confidently reject such arguments on the analytic grounds that the Axiom of Transitivity is necessarily true in virtue of the meanings of the words “all-things-considered better than”.

In light of the foregoing, let us quickly revisit what appears to be going on in section I's initial Spectrum Argument. The First Standard View reflects an *additive-aggregationist* approach that seems relevant and significant for certain comparisons. That is, in comparing the first alternative with second, it seems appropriate to basically multiply the intensity of the pain times its duration, in determining which of the two alternatives is better, and this yields the plausible judgment that the first alternative (the slightly more intense pain of shorter duration) is better than the second (the slightly less intense pain of much longer duration). Similar additive-aggregationist reasoning seems appropriate in comparing the second alternative with the third, the third with the fourth, the fourth with the fifth, and so on. However, the Second Standard View reflects an *anti-additive-aggregationist* approach that seems relevant for *other* comparisons. In particular, in comparing the first alternative with the last, most people *don't* simply multiply the intensity of the pains times their durations. Rather, they judge that where the difference in intensity of the pain is such that the more intense pain of a given duration has a significantly adverse effect on one's life, while the less intense pain of much longer duration would have little adverse effect on one's life, then the former would be much worse than the latter, even though the *sum total* of pains as determined by their intensities times durations would be greater in the latter situation than the former. So, in essence, most people believe that one set of criteria is relevant and significant for assessing how bad the first alternative is in comparison with the second, but a *different* set of criteria is relevant and significant for assessing how bad the first alternative is in comparison with the last. This reflects an Essentially Comparative View for assessing outcomes and, as we have seen, such a view opens up the door to rejecting the Axiom of Transitivity.

In response to my Spectrum Arguments, some total utilitarians and economists would reject the anti-additive aggregationist reasoning of the Second Standard View, and just insist that as long as there are *enough* extra mosquito bites, the life involving 16 mosquito bites per month *is* worse than the life involving two years of excruciating torture and fifteen mosquito bites per month. But is such a view really *plausible*?

Here are three related cases where most people would oppose simple additive aggregation. Most people firmly believe that Derek Parfit's Repugnant Conclusion is, indeed, repugnant (Parfit, 1984: ch. 17). They believe that an outcome, A, of at least ten billion people, all with a *very* high quality of life, would be better than an outcome, Z, with a vast population all of whom have lives that are *barely* worth living, no matter *how many* people live in Z. Similarly, most firmly believe that an *incredibly* flourishing human life that lasted, say, a *million* years, would be *better* for the liver of that life than a mere oyster-like existence, no matter *how many* years one might live in an

oyster-like state.⁸ And likewise, most firmly believe that no matter *how many* people would each get *one* lick of a lollipop, it would be better for that not to occur, if it unavoidably involved an innocent person suffering unbearable agony for many years followed by a slow, lonely, miserable death.⁹

Notoriously, total utilitarians reject such claims. Insisting that *more* utility is better than *less* utility, they offer a number of sophisticated explanations for why our intuitions about such cases are not to be trusted. For the total utilitarian, then, no matter *how* small the amount of good may be in a life that is *barely* worth living, or in a moment of oyster-like existence, or *how* small the amount of pleasure may be from *one* lick of a lollipop, if only there are *enough* such lives, moments, or licks, eventually the *total* amount of good or pleasure *will* be greater, and then be *better*, than, *any* finite amount of good or pain that might be balanced off against it.

The utilitarian's position is admirably consistent, but it reminds one of Emerson's contention that "a foolish consistency is the hobgoblin of little minds, adored by little statesmen and philosophers and divines" (Emerson 1983). Few are willing to "bite the utilitarian's bullet" in such cases, and I believe they are right not to do so. In evaluating outcomes, we don't simply care about how *much* utility obtains, we *also* care about how that utility is *distributed* and the *impact* that the distribution has on people's lives.

4. NEUTRALITY AND DOMINANCE PRINCIPLES

It is common for philosophers and others to assume that in certain contexts, morality requires us to be *neutral* with respect to people, places, and times. So, for example, setting aside the special obligations that one may have towards people with whom one stands in certain special relations —such as one's family, friends, students, patients, and so on— it is thought that, other things equal, if one could save one person or five, it would be better to save the five whether the five were (a) black or white, rich or poor, Hindu or non-Hindu, men or women, European or African, and so on (neutrality with respect to people), (b) close or far (neutrality with respect to space), or (c) living in the present, the near future, or the distant future (neutrality with respect to time —we'd also think it wouldn't matter if the five were living in the *past* if, contrary to fact, we could save people who were living in the past).

8. I discuss this kind of example, which I call the *Single Life Repugnant Conclusion*, in Temkin, 2012: ch. 4. The Single Life Repugnant Conclusion was originally presented by J. M. E. McTaggart (1921: vol. 2, 452-3).

9. My *Lollipops for Life* case is presented in Temkin, 2012: ch. 2. That case serves as the inspiration for the book's cover art.

Now I am aware that certain prevalent theories of modern physics discuss the space/time continuum in a way that suggests that space and time are not really distinguishable, so that however we treat space we should also treat time, and vice versa. But despite this, I have my doubts whether space and time should, in fact, be treated the same *normatively*. Consider, for example, the following thought experiment.

Suppose I learn that our civilization will live in our galaxy another 1000 years, and then die out. I also learn that in a distant galaxy another advanced civilization will exist for the same 1000 years and then die out, and that this is also so in a third distant galaxy, and a fourth distant galaxy. I find this all quite interesting. It is somewhat pleasing to me to learn that there are, in fact, advanced civilizations living in galaxies far away. But suppose I also learn that *beyond* the fourth galaxy there is nothing but cold, empty, space. This, too, I find interesting, but I must confess that learning that fact doesn't bother me at all. Indeed, if someone said that events beyond the fourth galaxy were about to unfold which would make those distant reaches inhospitable to life forms in perpetuity, I wouldn't think it important for our civilization to make significant sacrifices, if it could, to prevent that from happening.

Suppose, on the other hand, I vary the story a bit. As before, I learn that civilization in our galaxy will die out in 1000 years; but I learn that after ours dies out another advanced civilization will arise and persist for 1000 years in a second galaxy, and that this will happen again a third and fourth time. But I also learn that after the fourth civilization dies out there will be nothing but cold, empty, space, *forever*. For some reason, *that* knowledge would bother me a *lot*. Indeed, if I learned that events were about to unfold which would make the universe uninhabitable for any life forms 4000 years from now, unless our civilization made significant sacrifices to prevent that from happening, I would feel quite strongly that we should do so, and I would feel that way even if I knew that *our* civilization was going to die out in 1000 years no matter what we did.

My views here may ultimately be indefensible, but I don't think they are idiosyncratic. They reveal an asymmetry in my thinking about space and time. I think it very important that many periods of *time* are filled with flourishing sentient beings. I think it much less important that many areas of *space* are filled with flourishing beings.

There is much more to be said about this suggested asymmetry between space and time, but I shall not pursue this here. Instead, let me turn to another set of views that might be held regarding space, time, and people. At first blush, I think most people would readily accept the following three dominance principles: (1) if outcome A is better than outcome B at *every point in space*, then A is better than B; (2) if outcome A is better than out-

come B at *every moment in time*, then A is better than B; and (3) if outcome A and outcome B involve the very same people, and A is better than B for *every* person, then A is better than B.

1, 2, and 3 are exceedingly weak Pareto-like principles. According to the Pareto Principle, if two outcomes involve the same people, and the first outcome is better for at least one person and at least as good for everyone else, then the first outcome must be better than the second. 1 and 2 apply similar reasoning to the domains of space and time, respectively, as to the domain of people. In addition, 1, 2, and 3 require that the first outcome be better than the second at *every* point in space, at *every* moment in time, or for *every* person, respectively.

Given the widespread appeal of the Pareto Principle, the fact that the dominance principles noted above are much weaker—and are therefore even *more* plausible— than the standard Pareto Principle, and the common assumption that we should be neutral with respect to people, places, and times, I think it is fair to assume that most people would find *each* of the three dominance principles intuitively appealing. Indeed, I suspect that many people would think that each of the dominance principles is “obviously” true. Yet, it is easy to see that however intuitively appealing the three dominance principles may be, at least one of them must be rejected.

Consider Diagram One.

Day 1 P ₁ Hell Day 2 P ₁ Heaven; P ₂ , P ₃ Hell Day 3 P ₁₋₃ Heaven; P ₄₋₉ Hell Day 4 P ₁₋₉ Heaven; P ₁₀₋₂₇ Hell : : W ₁	Day 1 P ₁ Heaven Day 2 P ₁ Hell; P ₂ , P ₃ Heaven Day 3 P ₁₋₃ Hell; P ₄₋₉ Heaven Day 4 P ₁₋₉ Hell; P ₁₀₋₂₇ Heaven : : W ₂
--	--

Diagram one

Diagram One represents two possible worlds God is thinking of instantiating, W₁ and W₂. In W₁, there will be a single person, P₁, who will exist on Day 1, and he will be in Hell. We don't have to think that P₁'s life will be infinitely bad, we just have to think that it will be very bad. During the course of that day, it would be much better for P₁ if he were not alive. On Day 2, P₁ moves to Heaven, where it will be very good for P₁ that he is alive. For simplicity, let us assume that each day in Heaven would be as good for the person experienc-

ing it as a day in Hell would be bad for a person experiencing it, so that on balance the net value of a life with an equal number of days in Heaven and in Hell would be zero. Unfortunately, on Day 2 two new people, P_2 and P_3 are created and put in Hell. On Day 3, each of P_1 - P_3 are in Heaven, but six new people P_4 - P_9 are in Hell. On Day 4 each of P_1 - P_9 are in Heaven, but 18 new people are created in Hell. And so on.

W_2 is just like W_1 except in reverse. In W_2 , P_1 will again exist on Day 1, but this time he will start in Heaven. On Day 2, P_1 moves to Hell, but two new people, P_2 and P_3 are created and put in Heaven. On Day 3, each of P_1 - P_3 are in Hell, but six new people P_4 - P_9 are in Heaven. On Day 4 each of P_1 - P_9 are in Hell, but 18 new people are created in Heaven. And so on.

How do W_1 and W_2 compare in terms of goodness? Which, if either, is the better outcome, all things considered? If one looks at the two outcomes day by day, it may seem clear that W_2 is better than W_1 . After all, on Day 1, there would be one person in Hell in W_1 and one person in Heaven in W_2 . So, on Day 1, W_1 is clearly worse than W_2 . Similarly, on Day 2, W_1 would have *one* person in Heaven, but *two* people in Hell, whereas W_2 would have *one* person in Hell, but *two* people in Heaven. Given our views about neutrality with respect to people, it seems clear that it is worse for there to be twice as many people in Hell as in Heaven, than it is for there to be twice as many people in Heaven as in Hell, so W_1 is worse than W_2 on Day 2. Similarly, on Day 3, W_1 , where there are three people in Heaven but six people in Hell, will be worse than W_2 , where there are three people in Hell, but six people in Heaven. And so on. The point is that on Day 1, W_1 is worse than W_2 , and that on *each* day after that W_1 is worse than W_2 , since, on each day after Day 1, there will always be twice as many people in Hell as in Heaven in W_1 , while there will always be twice as many people in Heaven as in Hell in W_2 . Thus, comparing W_1 and W_2 day by day, or moment by moment, the dominance principle with respect to *time* would entail that W_2 is better than W_1 .

Is W_2 better than W_1 ? I find that *very* hard to believe. Suppose we compare the two outcomes not moment by moment, but person by person. In W_1 , each person spends exactly *one* day in Hell, and the rest of *eternity* in Heaven. In W_2 , each person spends exactly *one* day in Heaven, and the rest of *eternity* in Hell. I know which of these worlds I would want for myself, a loved one, or anyone else who was not pure evil! I would want W_1 , and I would want it because it would be better for *each* person who ever lived. Notice, since in this example we are assuming that the very same people would live in each world, and we know that each of them would be better off in W_1 than W_2 (indeed *vastly* so, since it is *much* better to spend only one day in Hell and the rest of eternity in Heaven, than to spend only one day in Heaven and the rest of eternity in Hell), then the dominance principle with respect to *people* would entail that W_1 is better than W_2 .

In this example, we see that two intuitively plausible and seemingly “obvious” dominance principles are in fact incompatible. In this case, at least, we must choose between the dominance principle with respect to time and the dominance principle with respect to people. As I have already made clear, I know how *I* would choose in this case. I think W_1 is clearly and unequivocally better than W_2 .

Notice, if one adopted a purely *impersonal* view of morality, according to which it didn't matter how any particular sentient beings fared, or how benefits or burdens were distributed within or between lives, but it *only* mattered how *many* benefits or burdens obtained in an outcome, then it might be plausible to maintain that W_2 is better than W_1 , in accordance with the dominance principle with respect to time, or, alternatively, that W_1 and W_2 were *equally* good, since each would ultimately involve an infinite number of days lived in both Heaven and Hell of the same orders of infinity. But my own view is that one lesson to be learned from Diagram One is that in assessing the goodness of outcomes we should *not* merely focus on the impersonal questions of how *much* well-being there is in the two outcomes, or how *many* benefits and burdens obtain in total. Rather, in some cases, at least, we must focus on the question of how the well-being or benefits and burdens are *distributed*, and, in particular, on how the sentient beings are *affected* for better or worse in those outcomes.

5. ESSENTIALLY COMPARATIVE IDEALS

I claimed earlier that a number of ideals people attach great value to have an Essentially Comparative structure, including the Pareto Principle, the most plausible versions of Maximin and Utility, and the Narrow Person-Affecting Principle. I defend this claim in Temkin (2012: ch. 12) for each of the ideals in question, but for the purposes of this article let me just focus on the Narrow Person-Affecting Principle.

In any choice situation between possible outcomes, let us call those people who do exist, or have existed, or will exist in each of the outcomes independently of one's choices, *independently existing people*. By contrast, let us call those people whose existence in one or more possible outcomes depends on the choices one makes in bringing about an outcome, *dependently existing people*. Bearing these distinctions in mind, we can now state the Narrow Person-Affecting View.

The Narrow Person-Affecting View: In assessing possible outcomes, one should (1) focus on the status of independently existing people, with the aim of wanting them to be as well off as possible, and (2) ignore the status of dependently existing people, except that one wants to avoid harming

them as much as possible. Regarding clause 2, a dependently existing person is harmed only if there is at least one available alternative outcome in which that very same person exists and is better off, and the size of the harm will be a function of the extent to which that person would have been better off in the available alternative outcome in which he exists and is best off.¹⁰

As stated, the Narrow Person-Affecting View reflects an important extension of Jan Narveson's claim that "Morality has to do with how we treat whatever people there are.... [We] do not ... think that happiness is impersonally good. We are in favor of making people happy, but neutral about making happy people" (Narveson 1973: 73 and 80). Specifically, the first clause reflects the view that we are neutral about making people exist, while the second clause reflects the important qualification that if we *are* going to make a particular person exist, her interests have to count the same way as every other existing person's, in that we must equally seek to make that person, like every other existing person, as well off as possible.

Now, in fact, that there are lots of ways in which the Narrow Person-Affecting View needs to be qualified and limited in scope, which I won't go into here (Temkin 2012: ch. 12.3). Nevertheless, when properly interpreted, the Narrow Person-Affecting View reflects a deeply plausible and widely-accepted view for a certain range of cases.

To illustrate the Narrow Person-Affecting View, it will be useful to consider a range of cases to which it might be applied, and to contrast it with some other principles that might be appealed to in assessing outcomes: the *Impersonal Total View*, the *Impersonal Average View*, and the *Wide Person-Affecting View*. Roughly, we might say that the *Narrow Person-Affecting View* assesses outcomes by considering how the *particular* people in those outcomes fare, relative to how they fare in any available alternative outcomes (here, and below, "people" refers to any sentient beings). The aim is to make sure that each *particular* person who does, or will, exist independently of our choices, or who will exist as a result of our choices, fares as well as possible. In contrast, the *Wide Person-Affecting View* assesses outcomes by considering how the *people* in those outcomes fare, but it is not concerned with how any *particular* people fare in one outcome relative to how those very

10. Derek Parfit presented a position which he also called a Narrow Person-Affecting View in Parfit, 1984: ch. 18. The view as I present it here is different in important respects than Parfit's, but I have retained the name Parfit uses, because I think the view I have described reflects a fundamental approach to assessing outcomes that is best described as a Narrow Person-Affecting View. I believe that my version of the Narrow Person-Affecting View is more plausible than Parfit's original version, and in conversation Parfit has indicated that he agrees.

same people might fare in any available outcomes.¹¹ A precise characterization of the Wide Person-Affecting View is elusive, but one natural and plausible way of interpreting it implies, among other things, that if the people in one outcome, A, are all better off than the people in another outcome, B, whether or not they are the *same* people or there are the same number of people, then A is better than B; if, for each distinct person in B, there is corresponding distinct person in A, at least one of whom is better off and the rest of whom are at least as well off, then A is better than B as long as anyone else existing in A has a life that is (sufficiently) worth living; and if A and B have the same number of people, and for each person in B there is a corresponding person in A who is equally well off, and vice versa, then A and B are equally good. Finally, the Impersonal Total and Average Views imply that regardless of whether or not they have the same people or the same number of people, one outcome will be better than (equal to) another if and only if the one outcome has a higher (the same) total or average amount of utility or well-being, respectively.

Consider Diagram Two.

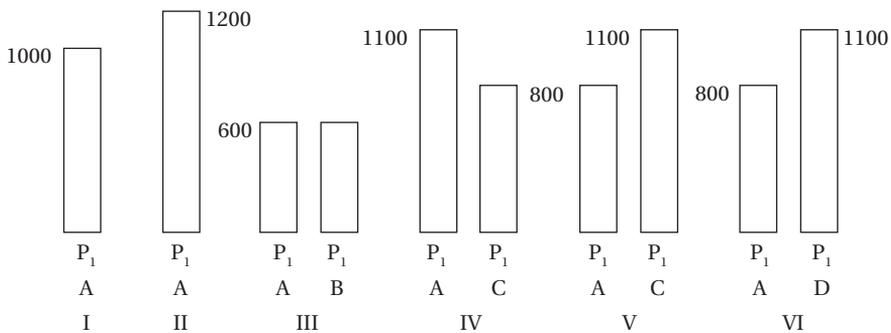


Diagram two

In I, there is a large population, A, say of 10 billion people, on a given planet, P₁, all of whose members are at level 1000. Assume that I is the initial outcome, and that the A people are thinking about transforming their outcome into one represented by II. In II, those very same people all exist and are

11. The notion of a Wide Person-Affecting View was introduced by Parfit, 1984: ch. 18. Unfortunately, as Parfit originally presented the position, he combined two elements which are best kept distinct. The first reflects the view that in assessing outcomes we want to assess them in terms of the extent to which the people (sentient beings) in those outcomes are affected for better or worse. The second concerns the very distinct question of whether causing someone to exist benefits that person. I use the notion of a Wide Person-Affecting View to reflect the first element only. Parfit now shares my view (Temkin, 2012: note 41, section 12.4).

better off, at level 1200. II would be judged a better outcome than I on all of the different approaches for assessing alternatives. Specifically, II is better than I on the Impersonal Total and Average Views, since the total and average amounts of wellbeing are greater in II than in I, on the Wide Person-Affecting View, since it is better for people, as everyone in II is better off than everyone in I, and on the Narrow Person-Affecting View, since it is better for the particular, independently existing A people who exist in both outcomes.

Suppose, instead, that the A people could transform I into an outcome like III. In III, the A people have all been lowered to level 600, but a new population of 10 billion people, B, would also come to exist at level 600 on a second planet, P_2 . In this scenario, III would be ranked better than I on the Impersonal Total View, since the total wellbeing would be greater in III than in I. But III would be ranked worse than I on the Impersonal Average View, since the average level of wellbeing would be less than in III than in II. III would also be ranked worse than I on the Wide Person-Affecting View, since the people in I are better off than the people in II. Finally, III would also be ranked worse than I on the Narrow Person-Affecting View, as the independently existing A people are better off in I than in III, and the principal aim of the Narrow Person-Affecting View is to make the particular existing people as well off as possible (making people happy) rather than to add more people to an already large and well-off population (making happy people).

While total utilitarians would rank III better than I, if outcome I were one's starting point, many people, and perhaps most, would rank I better than III, and they might do so on any combination of the grounds suggested.

Suppose next that the people in I could bring about IV. IV involves a new group of 10 billion people, C, living on a different planet, P_3 . Unfortunately, the conditions on P_3 are not quite as favorable as those on P_1 , so the C people would only be at level 800. But we may presume that level 800 is still quite high, so that everyone on P_3 would have lives well worth living. In addition, there might be resources on P_3 which could be used in trades with those on P_1 , so that everyone in P_1 would be raised up to level 1100.

IV would be worse than I on the Impersonal Average View. Many find this hard to believe. If there is an objection to IV, it would seem to rest on the fact that IV involves *inequality* while I is perfectly equal, not on the fact that the *average* level of well-being is lower in IV than in I. After all, IV is better off than I for *everyone* who lives in I, and *in addition* IV involves a very large group of people all of whom have lives that are well worth living.¹²

12. The strongest arguments against the average view involve alternatives where people's lives are well below the level at which life ceases to be worth living. Surely, one wouldn't improve an outcome where billions of people were living in the worst hell imaginable in *any* respect, merely by adding billions of more people whose lives were almost, but not quite, as badly off.

On reflection, I believe most people would judge IV better than I, and this would be supported by the Impersonal Total View —since the total wellbeing is greater in IV than in I— by the Wide Person-Affecting View —since IV is better for people than I, as for each person in I there is a corresponding person in IV who is even better off, and any additional people have lives that are well worth living— and by the Narrow Person-Affecting View, since the particular independently existing A people are better off in IV (being at level 1100) than in I (being at level 1000).

Next, suppose that the people in outcome I could bring about either IV or V. In V, the A people have to make extra sacrifices to enable the C people to live on a different, more hospitable, fourth planet P_4 . The result would be that the C people would be at level 1100, but the A people would only be at level 800. Interestingly, as alternatives to I, IV and V would likely be regarded as equally good on all four of the principles we have been discussing. IV and V are equally good on the Impersonal Total and Average Views, as they are equally good in terms of total and average wellbeing. They are equally good on the Wide Person-Affecting View, since in terms of how people in those outcomes fare (rather than in terms of how the *particular* people fare in one outcome rather than another), they are equally good for people. Finally, they are equally good on the Narrow Person-Affecting View, since on that view one doesn't *have* to bring about the dependently existing C group, but *if* one is going to bring a particular group into existence —and, by hypothesis, the very same C people would be brought into existence in both IV and V— then their interests have to be given the *same* weight as those of the independently existing people, A. Hence, on the Narrow Person-Affecting View, there would be nothing to choose between outcome IV, where the independently existing A people would be at level 1100 and the dependently existing C people would be at level 800, and outcome V, where the independently existing A people would be at level 800 and the dependently existing C people would be at level 1100.

Finally, suppose that the option facing those in I is not IV or V, but IV or VI. Here, the option is between populating planet P_3 with 10 billion people, the C people, who would all be well off, but “only” at level 800, but where this would enable the A people to raise their level from 1000 to 1100, or populating a more hospitable but more distant planet P_4 , with an entirely *different* group of 10 billion people, the D people, but where the cost of populating the more distant planet would be to *lower* the level of the A people from 1000 to 800. IV and VI would be *equally* good on both the Impersonal Total and Average Views, since the total and average levels of wellbeing are equal

But, of course, the addition of all those extra people living hellish lives would raise the average level, even if only by a small amount. For further discussion of this kind of case, which Parfit called *Hell Three*, and other reasons to be skeptical of average views, see Parfit, 1984: , 422; Temkin, 2012: section 10.4; Temkin, 1993: section 7.5.

in both outcomes. Likewise, IV and VI, would be equally good on the Wide Person-Affecting View, since, overall, people fare equally well in both outcomes. However, importantly, if one's initial starting place was I, then IV would be decidedly *better* than VI on the Narrow Person-Affecting View. This is because, insofar as we are concerned with "making people happy, rather than making happy people", IV is a clear improvement, while VI is a clear worsening of the outcome.

That is, on the Narrow Person-Affecting View, IV is *better* for the independently existing A people (they are at level 1100 rather than level 1000), and it in *no way harms* the dependently existing C people, since their lives are well worth living, and, in this choice situation, there is no available alternative in which they would be better off. VI, on the other hand, is clearly *worse* for the independently existing A people (they are at level 800 rather than level 1000), and this worsening of the outcome cannot be made up for by the *neutral* factor of adding extra "happy" D people.

Let me acknowledge that the Narrow Person-Affecting View is not plausible in cases like Parfit's *Non-Identity Problem* (Parfit 1984: ch. 16) As stated, it is also implausible in a host of other cases, many of which will readily occur to the reader. However, despite this, I believe that the Narrow Person-Affecting View is plausible, and relevant and significant for comparing outcomes in a large range of cases, including those just discussed. Thus, in considering cases like those represented in Diagram Two, I believe that many people would judge that if one's initial outcome were like I, then II would be better than I, III would be worse than II, IV would be better than I, IV and V would be equally good, and IV would be better than VI, and I believe that many would base their judgments partly, if not wholly, on narrow person-affecting grounds (or a position very much like it in spirit if not exact detail).

As should be clear, the Narrow Person-Affecting View is an Essentially Comparative Ideal. On such a view, assessing how good an outcome is depends not solely on its *internal* features, as is the case on the Internal Aspects View, but on whether the particular people in that outcome exist in available alternative outcomes, and if so, on how they fare in the available alternatives.

Assuming that there would be no morally relevant differences between the different people in my examples *other than* narrow person-affecting considerations, on the Internal Aspect View IV, V, and VI would be *equally* good, since their internal features are identical, except for which *particular* people exist in which outcomes and which particular levels they are at. Hence, in accordance with the Principle of Like Comparability for Equivalents, on the Internal Aspects View, however *one* of them compared with some other alternative, that is how *each* of them would compare with that alternative, and this would be so regardless of whether or not any other out-

comes were available. But, as we have seen, in accordance with the Narrow Person-Affecting View, many people would judge IV as *better* than I, if outcome I was the initial starting point and those were the only alternatives, but they would judge V as *worse* than I, if outcome I was the initial starting point and those were the only alternatives. Similarly, in accordance with the Narrow Person-Affecting View, many people would judge IV as *equally as good* as V, if those were the only alternatives, and V as *equally as good* as VI, if those were the only alternatives, but, contrary to both the Principle of Like Comparability for Equivalents and the *Axiom of Transitivity for Equally as Good As* (each of which is entailed by the Internal Aspects View), they would deny that VI is equally as good as IV. Likewise, in accordance with the Narrow Person-Affecting View, and contrary to the Axiom of Transitivity for Better Than, it is plausible to contend that if outcome I were one's initial starting point, then IV would be *better* than I if those were the only alternatives, and I would be better than V if those were the only alternatives, but IV would *not* be better than V if those were the only alternatives.

Could we abandon the Narrow Person-Affecting View and simply adopt Impartial Views or the Wide Person-Affecting View instead? Not without abandoning a view that underlies many judgments people make in assessing outcomes. And not easily. To buttress this claim, let us consider two further cases, of a different sort, the first of which is exemplified by Diagram Three.

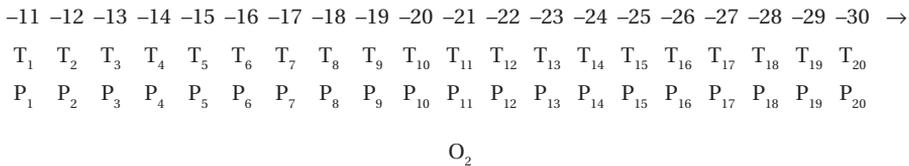
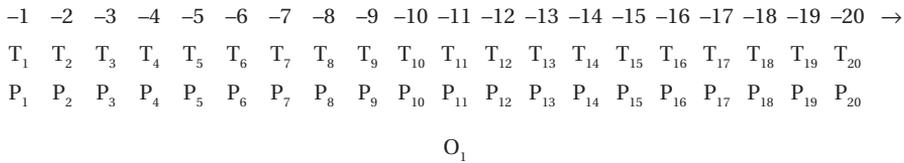


Diagram three

Suppose that one of two outcomes was going to come about. In O₁, there would be one person living on planet one, P₁, at time one, T₁, and that person would be at level -1, which is below the level at which life ceases to be worth living. It would be better for that person if he or she never existed. There would also be one person living on planet two, P₂, at time two, T₂, and that person would be

even worse off at level -2 . There would be a third person living on planet three, P_3 , at time three, T_3 , and that person would be even worse off at level -3 , and so on. Hence, there would be an infinite number of people living on different planets and at different times, and each person, after the first, would be worse off than those that preceded him or her. In addition, let us assume there would be no other morally relevant factors or events obtaining in W_1 .

In the second outcome, O_2 , there would again be one person living on planet one, P_1 , at time one, T_1 , but this time the person would be at level -11 . There would also be one person living on planet two, P_2 , at time two, T_2 , and that person would be even worse off at level -12 . There would be a third person living on planet three, P_3 , at time three, T_3 , and that person would be even worse off at level -13 , and so on. As before, there would be an infinite number of people living on different planets and at different times, and each person, after the first, would be worse off than those that preceded him or her, and there would be no other morally relevant factors or events obtaining in O_2 . Finally, for any level $-n$, it is worse for someone to be at level $-(n - 10)$, than to be at level $-n$.

How do O_1 and O_2 compare? As described, there might be some reasons associated with how we think about cases involving infinity, for claiming that O_1 and O_2 were equally good. On the other hand, I think there would also be powerful reasons for thinking that O_1 was *better* than O_2 . If we, or God, had to choose which of the two outcomes to produce, or we learned that one of the two outcomes was going to be instantiated, at first blush it seems that we should produce or hope that it is O_1 rather than O_2 . Other things equal, it seems we should be confident that O_1 would be *at least at good as* (and probably *better than*) O_2 .

The preceding ranking of O_1 and O_2 would be supported by both Impersonal Principles, as well as any plausible Wide Person-Affecting View. Insofar as one merely focuses on the *impersonal* value in each outcome, or on how *people* fare in each outcome without regard to how any *particular* people fare, then it seems clear that O_1 is *at least at good as* O_2 . Moreover, I think O_1 *would* be at least at good as O_2 if completely *different* people lived in O_1 than in O_2 , or if anyone who lived in *both* outcomes, lived on the *same* corresponding planet and at the *same* corresponding time in both outcomes. Thus, if a given person, John, lived in both outcomes, then *whatever* planet P_n and time T_n that he occupied in O_1 , he would also occupy P_n and T_n in O_2 .

Suppose, however, that I now tell a different story regarding the members of O_1 and O_2 . Suppose it is true that *every* person who would exist in O_2 , if O_2 obtained, would *also* exist in O_1 , if O_1 obtained, but that *each* of them would be *worse off* in O_1 than in O_2 . Specifically, let us assume that the *very same* individual, I_1 , who would exist at T_1 in O_2 , would exist at T_{21} in O_1 , that the *very same* individual, I_2 , who would exist at T_2 in O_2 , would exist at T_{22} in O_1 , that the *very same* individual, I_3 , who would exist at T_3 in O_2 , would exist at T_{23} in O_1 , and

so on. It would then be the case that *every single person* who would exist in O_2 would also exist in O_1 and would be ten units worse off, where, as before, for *any* level $-n$, it is worse for someone to be at level $-(n - 10)$, than to be at level $-n$.

Given that scenario, it seems clear that if we, or God, had to choose which of the two outcomes to produce, or we learned that one of the two outcomes was going to obtain, we should produce or hope that it is O_2 rather than O_1 ! O_1 is worse than O_2 for *every person* who lives in O_2 , and, *in addition*, there are 20 *different* individuals who exist in O_1 but not in O_2 (those who would be living at times T_1 through T_{20} in O_1), whose lives are below the zero level—they would rationally prefer that they had never been born. Surely, if we were aiming to choose the *better* outcome, and we knew that we or our loved ones might actually be occupants of one of the two worlds, we would choose O_2 , and we would make a similar choice on behalf of any strangers who were not pure evil.

It seems clear, then, that our judgments about how outcomes like O_1 and O_2 compare would not, and *should* not, be influenced solely by *impersonal* or *wide* person-affecting considerations. In some cases, how the *particular* people are affected for better or worse depending on the alternatives is *rightly* relevant to our assessment, as is implied by the Narrow Person-Affecting View. Thus, in some cases at least, cross-world identification of particular individuals is both relevant and necessary for accurately comparing outcomes, as is permitted on the Essentially Comparative View of ideals, but is prohibited by the Internal Aspects View.

Let us apply the preceding reasoning to a final case, represented by Diagram Four.

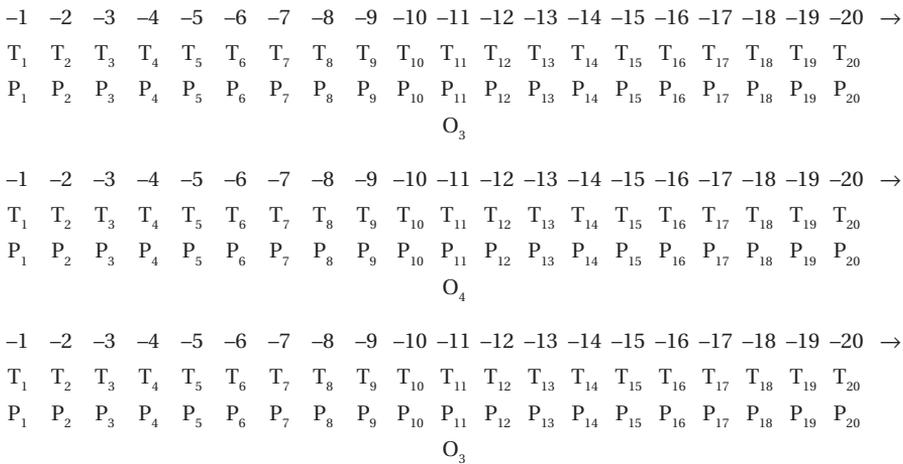


Diagram four

O_3 , O_4 , and O_5 are just like O_1 in Diagram Three. In each outcome there is one person on P_1 at T_1 at level -1 , a second person on P_2 at T_2 at level -2 , a third person on P_3 at T_3 at level -3 , and so on. If one asked how O_3 , O_4 , and O_5 compared, it would be natural to assume that they were all *equally good*, all things considered. And if there were *different* people in O_3 , O_4 , and O_5 , then it seems clear that they *would* all be equally good.

Suppose, then, we make the assumption that the people in O_3 would be different people than those in O_4 , and similarly that the people in O_3 would be different people than those in O_5 . In that case, there would be no narrow person-affecting considerations that were relevant for comparing O_3 with O_4 , or for comparing O_3 with O_5 , and there would be good grounds for judging that O_3 and O_4 were *equally good*, and similarly that O_3 and O_5 were *equally good*. Does it follow from this that O_4 and O_5 must be equally good, as it *must* if the Internal Aspects View is correct, since such a view entails both the Principle of Like Comparability for Equivalentents and the transitivity of the “equally as good as” relation? It does not! Because consistent with the forgoing relations between O_3 and O_4 , and O_3 and O_5 , O_5 may stand in a *different* relation to O_4 , one that is similar to the relation in which O_2 stood to O_1 in Diagram Three.

After all, even it is *true* that the people in O_3 are *different* from the people in *both* O_4 and O_5 , it *doesn't* follow from that that the people in O_5 are different from the people in O_4 . They may not be! Suppose, then, that the person who would occupy P_1 and T_1 and be at level -1 in O_5 , would occupy P_{11} and T_{11} and be at level -11 in O_4 , the person who would occupy P_2 and T_2 and be at level -2 in O_5 , would occupy P_{12} and T_{12} and be at level -12 in O_4 , the person who would occupy P_3 and T_3 and be at level -3 in O_5 , would occupy P_{13} and T_{13} and be at level -13 in O_4 , and so on. It would then be the case that *everyone* who exists in O_5 *also* exists in O_4 and is ten units worse off, and that, *in addition*, there would be 10 *different* individuals who exist in O_4 , but not in O_5 , whose lives would be below the zero level and who would rationally wish that they had never been born. In this case, as above, it seems clear that O_4 would be a *worse* outcome than O_5 , and mainly in virtue of narrow person-affecting considerations.

We see, then, that in accordance with the Essentially Comparative View, a factor that is relevant and significant for comparing O_4 and O_5 —specifically, the fact that everyone who exists in O_5 also exists in O_4 where he or she is worse off—is not relevant or significant for comparing O_3 with O_4 , or O_3 with O_5 . This explains how it can be the case that in terms of *all* of the factors that are relevant and significant for making *each* comparison, O_3 and O_4 might be equally good, and O_3 and O_5 might be equally good, but O_4 and O_5 might not be equally good.

More generally, as we have seen, once we accept an Essentially Comparative View of ideals, as it seems we must if we are to account for the judgments to which many are committed regarding Diagrams Two, Three, and Four, then there is no reason to expect the “all-things-considered better than” or “all-things-considered equally as good as” relations to be transitive, or, alternatively, no reason to think that such relations even *apply* to various alternatives we may have expected them to for purposes of practical reasoning.¹³

6. CONCLUDING REMARK

As promised at the beginning, this article barely scratches the surface of some of the issues raised in *Rethinking the Good*. Moreover, the further one explores such issues, the more one realizes how this domain is fraught with complications, unresolved difficulties, and impossibility results whose premises are exceedingly difficult to abandon. The book seriously challenges us to rethink our understanding of the good, moral ideals, and the nature of practical reasoning in many ways that have deep practical and theoretical implications. But beyond that, I’m afraid, it offers little guidance, and I have little sense, of where we go from here. I wish it were otherwise.

BIBLIOGRAPHY

- Broome, J., 1991: *Weighing Goods*, Oxford: Basil Blackwell.
 — 2004: *Weighing Lives*, Oxford: Oxford University Press.
 Emerson, R. W., 1983: “Self-Reliance”, in *Essays and Lectures*, ed. Joel Porte, New York Library of America: 259-82.
 McTaggart, J. M. E., 1921: *The Nature of Existence*, 2 vols., Cambridge: Cambridge University Press).
 Narveson, J., 1973: “Moral Problems of Population”, *The Monist* 57.
 Parfit, D., 1984: *Reasons and Persons*, Oxford: Oxford University Press.
 Rachels, S., 1993: “A Theory of Beneficence”, unpublished Philosophy, Politics and Economics thesis, Oxford University.
 — 1998: “Counterexamples to the Transitivity of Better Than”, *Australian Journal of Philosophy* 76: 71-83.

13. In my book, I discuss various ways of preserving the Axioms of Transitivity in the face of my arguments, which have the implication that there is no single set of alternatives that are being compared in the cases I discuss, or no single relation that the different alternatives are being compared in terms of, so that there is, strictly speaking, no *violation* of the axioms of transitivity in the cases I discuss, rather those axioms don’t even *apply* to the cases I consider. I suggest that even if such a move can be plausibly defended, it has significant practical and theoretical difficulties akin to those that would accompany the rejection of the Axioms of Transitivity (see Temkin, 2012: ch. 13).

- 2001: “A Set of Solutions to Parfit’s Problems”, *Noûs* 35: 214-238.
- 2004: “Repugnance or Intransitivity”, in *The Repugnant Conclusion: Essays on Population Ethics*, ed. J. Ryberg and T. Tannsjo: 163-86, Dordrecht: Kluwer Academic Publishers.
- Temkin, L., 1987: *Philosophy and Public Affairs* 16: 138-187.
- 1993: *Inequality*, New York: Oxford University Press.
- 2012: *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*, Oxford: Oxford University Press.
- 2014: “Intransitivity and the Internal Aspects View”, *Law, Ethics and Philosophy* 2: ??
- Wittgenstein, L., 1958: *Philosophical Investigations*, 2nd ed., ed. and translated by G. E. M. Anscombe, Oxford: Blackwell.

In Defense of the Internal Aspects View: Person-Affecting Reasons, Spectrum Arguments and Inconsistent Intuitions

ÓSCAR HORTA

University of Santiago de Compostela

Abstract

According to the Internal Aspects View, the value of different outcomes depends solely on the internal features possessed by each outcome and the internal relations between them. This paper defends the Internal Aspects View against Larry Temkin's defence of the Essentially Comparative View, according to which the value of different outcomes depends on what is the alternative outcome they are compared with. The paper discusses both person-affecting arguments and Spectrum Arguments. The paper does not defend a person-affecting view over an impersonal one, but it argues that although there are intuitive person-affecting principles that entail an Essentially Comparative View, the intuitions that support these principles can also be accommodated by other principles that are compatible with the Internal Aspects View. The paper also argues that the rejection of transitivity and the Internal Aspects View does not help us to solve the challenges presented by Spectrum Arguments. Despite this, the arguments presented by Temkin do succeed in showing that, unfortunately, our intuitions are chaotic and inconsistent. The paper argues that this has metaethical consequences that will be unwelcome by a moral realist such as Temkin, since they challenge the idea that our intuitions may track a moral reality existing independently of our preferences.

Keywords: betterness, Essentially Comparative View, Internal Aspects View, person-affecting reasons, Spectrum Arguments, transitivity.

1. INTRODUCTION

Larry Temkin's book *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning* (Temkin, 2012) is the most powerful challenge to our understanding of axiology and normative theory in contemporary philosophy to date. Calling it a classic or a masterpiece is an understatement: it

is a true milestone with which we enter a new stage in the study of value theory. Its importance can hardly be overestimated given how radically it will change axiology as well as normative ethics. The reason why this is so is that *Rethinking the Good* shows that a fundamental assumption in value theory that we used to consider axiomatic is not free from doubts and may be actually challenged. To understand why this is so, consider the following two opposing views about the factors according to which an outcome can be better than another:

The Internal Aspects View: the value of different outcomes, and whether a certain outcome is better or worse than another, depends solely on the internal features possessed by each outcome and the internal relations between them.

The Essentially Comparative View: the value of different outcomes, and whether a certain outcome is better or worse than another, depends on what is the alternative outcome they are compared with. A given outcome may have one value given one alternative, but a different value given another alternative. This can happen both when we compare outcomes regarding a certain ideal in particular or when we compare them as a whole, all things considered.

The Internal Aspects View is of course the traditional and more intuitive view concerning the value of outcomes. Value theory as we understand it today is based upon it. Yet Temkin has claimed that we should reject this view and accept instead an Essentially Comparative View. He has done this by presenting a set of arguments against the claim that betterness relations (“A is better than B”) must be transitive.¹ Those arguments show that in a number of cases it is just impossible for us to hold intuitive views regarding which outcomes are better than others unless we abandon the axiom of the transitivity of betterness. If we do this, however, that would mean that a very important part of what has been considered essential to practical reasoning until now will no longer apply.

Temkin’s arguments are so strong that from now on we will not be able to just take for granted, as most of us have done thus far, the Internal Aspects View. This does not mean, however, that we have to abandon this view, as Temkin argues. But it does mean that if we want to defend this view we will be now forced to respond to these arguments and not just presume the Internal Aspects View is true. In fact, this is what I intend to do in this paper.

1. See Temkin, 2012; see also 1987; 1996; 1997 and 2010. The other main proponent of this view has been Stuart Rachels, 1998; 2001 and 2004). See also Friedman, 2009. For criticisms of this view see Norcross, 1997; 1998 and 1999; Binmore and Voorhoeve, 2003; Voorhoeve, 2008; Broome (2004: 55-62).

I will argue that, powerful as Temkin's case for the Essentially Comparative View is, there are ways to resist it and to keep on maintaining the Internal Aspects View. I will claim that Temkin's arguments, instead, may drive us to doubt realist metaethics. In order to defend these claims, my argument will unfold as follows:

In section II, I will distinguish two main ways in which an Internal Aspects View can be challenged: by appealing to person-affecting essentially comparative principles such as the Narrow Person-Affecting Principle, and by appealing to Spectrum Arguments. I will claim that the kinds of reasons that are appealed to in each of these two lines of reasoning are essentially different ones. Due to this, whether one of them succeeds does not mean that the other will succeed too.

In section III, I will explain that while some principles such as the Narrow Person-Affecting Principle can lead to intransitive rankings of outcomes, the job this principle does in accommodating some common intuitions regarding person-affecting reasons can also be done by other principles as well. I will present three other principles that can do this: the *Time-Dependent Person-Affecting Principle*, the *Actuality-Dependent Person-Affecting Principle* and the *Identity-Dependent Person-Affecting Principle*. These principles entail the rejection of the transitivity of the “_ is better than _” relation. But they also seem to imply the asymmetry of that relation. I will claim that, despite this, when we examine the matter we discover that these principles do not really imply that betterness is not asymmetric or intransitive.² They only *appear* to have that implication. I will argue that, unlike the Narrow Person-Affecting Principle, they are not essentially comparative principles, and do not really question the Internal Aspects View. This does not mean that we should accept a person-affecting view instead of an impersonal one, but it does mean that those who accept it do not need to assume an Essentially Comparative View.

I will then argue in section IV that in Spectrum Arguments essentially comparative principles such as the Narrow Person-Affecting Principle do not play any role. I will then consider the puzzles we find in Spectrum Ar-

2. Note that nonasymmetry does not entail symmetry. The claim that betterness is an asymmetric relation means that if A is better than B, then B cannot be better than A. So the claim that betterness is not an asymmetric relation implies that if A is better than B, then B may or may not be better than A. But if betterness were a symmetric relation that would mean that if A is better than B, then B is better than A. Nonasymmetry is also different from antisymmetry. If betterness were an antisymmetric relation that would mean that if A is better than B, and B is better than A then A would be equal to B. Although accepting that betterness is intransitive entails rejecting it is transitive the opposite is not the case. Similarly, as Temkin (2012: 17) points out, transitivity of betterness may fail to apply in a certain set of alternative outcomes without that implying that intransitivity does apply. That would happen, for instance, if A were roughly equal to B, A were better than A' and A' were roughly equal to B. However, all the cases of nontransitivity we will see in this paper will also be cases of intransitivity.

guments, and will argue that they are not solved if we abandon the claim that betterness is transitive. We can see this when we consider what happens when we aim at global maximization instead of local maximization. This means Spectrum Arguments do not provide us with any conclusive reason to reject the Internal Aspects View.

Next, section V will conclude that in light of what has been argued in sections 3 and 4 the Internal Aspects View can be maintained despite the very strong objections Temkin has presented against it.

Finally, in section VI, I will argue that there is a conflict between the arguments against the Internal Aspects View presented by Temkin and his realist metaethical positions. Even though the problems presented by Temkin do not necessarily have metaethical implications, they still give us reasons to doubt moral realism. This is so because if moral reality exists we can only track it with our intuitions, and the arguments presented by Temkin show that our intuitions are chaotic and inconsistent. Temkin can only make his attack on the Internal Aspects View and the moral realist view he holds compatible by accepting a methodological approach that leaves room for inconsistency. But this clashes against some strong intuitions that many of us have towards consistency, as well as towards other metacriteria such as simplicity and transitivity. In fact, for many of us these intuitions can be stronger than those we have when we face pairwise comparisons of particular outcomes that contradict them. This is what will make us resist Temkin's arguments against the Internal Aspects View. But it will also drive us to doubt that those intuitions towards particular choices in pairwise comparisons of outcomes can be reliable, and thus to deny they can track any moral reality existing out there independently of us.

2. DIFFERENT WAYS TO QUESTION THE INTERNAL ASPECTS VIEW

In *Rethinking the Good* Temkin argues against the Internal Aspects View in different ways. I want to focus here in two of the strongest ones:

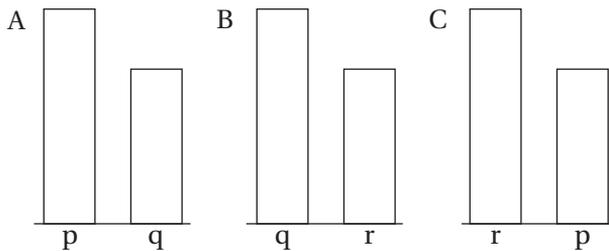
2.1. Appealing to Person-Affecting Reasons

Temkin tries to show that a consideration of person-affecting reasons is incompatible with the Internal Aspects View (2012: ch. 11 and 12).³ This would be so because according to these reasons whether a certain out-

3. In chapters 11 and 12 Temkin considers also other reasons apart from this one in favor of the Internal Aspects View that are closely related to the one I discuss here, but for lack of space I will not address them here.

come is better or worse than another depends on the relative situation of the individuals who are in them. If individuals in some outcome A are better off than they would be in outcome B, then A is better than B according to these principles. And if individuals in B are better off than they would be in C, then B is better than C. But this may be because the individuals that there are in A, B and C, and their relative situation in these outcomes, varies in such a way that makes it the case that the individuals in C are better off than they would be in A, so C would be better than A according to these principles. Imagine, for instance, that outcomes A, B and C are as follows:

Figure 1



Suppose now that we accept a principle such as the following one:

The Narrow Person-Affecting Principle: In assessing possible outcomes, one should (1) focus on the status of independently existing individuals, with the aim of wanting them to be as well off as possible, and (2) ignore the status of dependently existing individuals, except that one wants to avoid harming them as much as possible.⁴

According to this principle B would be better than A, since independently existing individuals would be better off in B. But for the same reason C would be better than B and A would be better than C. This means that if we accept this principle we have to reject transitivity and the Internal Aspects View. Due to this, we can consider it to be an essentially comparative principle, which we may define as follows.

Essentially comparative principles: a certain principle is essentially comparative if the factors for comparing two alternatives according to it may vary depending on the alternatives being compared, so an Essentially Comparative View necessarily is the case.

4. See Parfit (1984: 392); Temkin (2012: 417).

In other words, any principle whose acceptance entails accepting an Essentially Comparative View (as the Narrow Person-Affecting Principle) is an essentially comparative one.

2.2. Appealing to Spectrum Arguments

Another way to question the Internal Aspects View is by considering what happens in Spectrum Arguments (2012: ch. 2 and 5). In them we compare a number of different outcomes in which the values or disvalues that there are vary with respect to at least two dimensions. For instance, the outcomes may differ according to the intensity and the distribution through time or among different individuals of certain harms.⁵ Let us focus here, for the sake of simplicity, on a single-person Spectrum Argument Temkin discusses (2012: ch. 5). In it we compare outcome A, which is terrible torture for one year against outcome B, which is some only slightly milder torture for a much longer period. Then, we compare B against C, which is again some torture only slightly milder than the one in B, but for a much longer period. And we go on comparing each new outcome against a new one in which the torture is just a bit less painful but lasts for much longer. At the end, we reach outcome Z in which we experience some mild pain for a very long time—one which would be much longer than the total time we could ever live (Temkin suggests this could be the pain that a mosquito bite suffered each month would cause).⁶ Z is worse than another outcome Y in which we suffer for some pain which is only slightly higher but for a much shorter time.

In this spectrum, A seems clearly better than B, which seems clearly better than C, which seems better than D, etc., until we reach Y, which seems better than Z, which in turn, if transitivity applied, should be much worse than A. Yet Z seems intuitively better than A. Our intuitions in Spectrum Ar-

5. See Quinn, 1990.

6. This stipulation is unfortunate because it complicates unnecessarily the examination of the problem by introducing in it another dimension according to which our evaluation of different outcomes may vary: intermittence. By presenting Z as a situation in which one just suffers an extra mosquito bite for a month (in addition to other pains we may suffer, including other mosquito bites) we are not only considering intensity of pain and duration of pain, but also intermittence between different pains (this is not only so because a mosquito bite does not feel bad for a whole month, but also because mosquito bites do not itch continuously). Many of us would think it is worse to receive a more intense pain of some non trivial duration altogether in time than distributed in very short times spread in time. Consider, to see this, that a whole year of terrible torture seems to be worse than five seconds of torture each five years during a number of years equivalent to the number of seconds in a year. For the case to present the problem without including the problem of intermittence, the pain in Z should be felt continuously and uninterruptedly, although it would be an extremely mild pain.

guments therefore appear to entail that $A > B$; $B > C$; $C > D$; ... ; $Y > Z$; $Z > A$. In this way, if we accept what our intuitions tell us, transitivity does not seem to apply, and the Internal Aspects View fails.

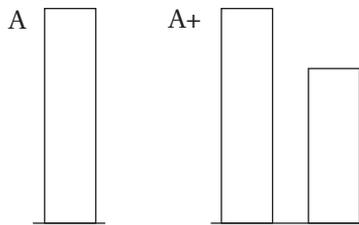
These two lines of reasoning, the one appealing to person-affecting reasons and the one appealing to Spectrum Arguments, are very different ones. Whether one of them succeeds is not dependent on whether the other one does. So in order to see if Temkin's case against the Internal Aspects View succeeds we will have to examine them separately.

3. THE APPEAL TO PERSON-AFFECTING PRINCIPLES

3.1. Principles that Can Entail Intransitivity: The Narrow Person-Affecting Principle

In order to explain why person-affecting principles challenge the Internal Aspects View Temkin considers what happens in the comparison that takes place between A and $A+$ in the Mere Addition Paradox (Parfit, 1984: ch. 142). As it is well known, in the Mere Addition Paradox A is an outcome in which a small group of individuals is enjoying a significantly high level of wellbeing. $A+$ is an outcome in which there are two groups: one of them is a group of individuals whose level of wellbeing is just like the one in A ; the other one is a group of individuals whose level of wellbeing is a bit lower than the one of those in A . The size of both groups is just like that of the group in A :

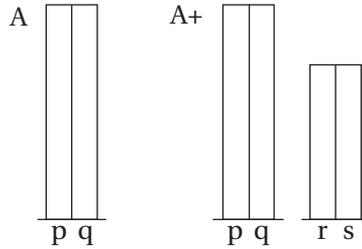
Figure 2



Is $A+$ worse than A ? Someone who considered this problem by taking into account exclusively impersonal reasons would reach always the same conclusion no matter the identity of those who are in A and $A+$. If she concluded that if $A+$ is worse than A , then that would mean that $A+$ has to be worse than A in all circumstances. And the same would happen if she concluded that $A+$ is better than A .

But those who accept person-affecting reasons can reach a different conclusion. The reason for this is that there are different ways in which we could move from A to A+. Consider first the one we can see in the next picture:

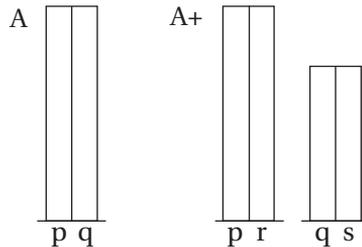
Figure 3



In this case, all the individuals who live in A are also present in A+ with the same level of wellbeing, but in A+ we add an extra group of individuals whose level of wellbeing is lower but still very good. According to a view that takes into account person-affecting reasons this would not make A+ worse than A: if anything, it would make it better.

But consider now this other way in which we may move from A to A+:

Figure 4



In this other case, half of those who are in A see their level of wellbeing reduced to the level of the worst off individuals in A+. And then, the population of both the better off and the worse off doubles. According to a view that takes into account person-affecting reasons and gives priority to the interests of those independently existing in both A and A+, this would make A+ worse than A. So A+ can be considered worse than A in some cases and not worse than A in others depending on the identity of those who are in

A+. ⁷This can be concluded, in particular, if one assumes a principle such as the Narrow Person-Affecting Principle. The rejection of the axiom of transitivity is an inescapable conclusion of the acceptance of this principle, at least as long as we face comparisons with different individuals.

This appears to be a powerful argument against the Internal Aspects View, although it depends on a view that is very controversial. It is not at all clear that we should accept a person-affecting view, and for those accepting a purely impersonal view Temkin's argument will have no force at all. In this paper I will remain neutral regarding whether we should accept an impersonal view or a person-affecting one. But I will argue that those who think that considering impersonal reasons alone has counter-intuitive implications need not accept essentially comparative principles such as the Narrow Person-Affecting Principle. There are different ways to defend the idea that person-affecting reasons must be considered. There are different person-affecting principles. The Narrow Person-Affecting Principle implies rejecting Transitivity, but other person-affecting principles lack this implication, even if they appear to possess it at first. Let me elaborate.

3.2. Person-affecting Principles that Seem to Imply Betterness Need Not Be Asymmetric

In order to examine this I need to point at an important distinction between different principles that appeal to person-affecting reasons. Some of these principles imply nontransitive or actually intransitive comparisons regarding the betterness of three or more outcomes, but despite this it is clear that according to them betterness is always asymmetric. The Narrow Person-Affecting principle is an example of these principles. If a certain outcome A is better than another one B, then according to this principle it is not possible that B is better than A according to it. This seems very intuitive. But there are other principles that appear to imply that betterness need not be asymmetric. ⁸ I will consider here three different principles that can imply this. There are other possible principles which can also fall within this class of principles, but for simplicity I will focus on these three principles that I think many of those defending a person-affecting view can find intuitive. In particular, I think many accept

7. See Temkin, 2012: ch. 11 and 12. The initial trigger for the development of this idea was Parfit's claim that the inequality occurring in A+ cannot be bad because it is produced by a mere addition of extra individuals (1984: 425). See on this claim also Temkin, 1997: 304. For criticisms see Dancy, 2005; Weber, 2007.

8. They also seem to imply, in a similar vein, that the “_ is at least as good as _” relation need not be antisymmetric.

The Actuality-Dependent Person-Affecting Principle: in assessing possible outcomes, one should (1) focus on the status of those who exist in the actual world, with the aim of wanting them to be as well off as possible, and (2) ignore the status of those who do not exist in the actual world, except that one wants to avoid harming them as much as possible.

I think many people also accept

The Time-Dependent Person-Affecting Principle: in assessing possible outcomes that have not occurred yet, one should (1) focus on those who will exist in the outcome that will occur first, with the aim of wanting them to be as well off as possible, and (2) ignore the status of those who will not exist in the outcome that will occur first, except that one wants to avoid harming them as much as possible.

And at least some philosophers also accept:

The Identity-Dependent Person-Affecting Principle: in assessing possible outcomes, one should (1) focus on the status of those whose identity is already determined, with the aim of wanting them to be as well off as possible, and (2) ignore the status of those whose identity is not determined yet, except that one wants to avoid harming them as much as possible.

I shall henceforth refer to these principles as “the three alternative person affecting principles” or simply “the three other principles”.

Now, to see these principles in action consider the way in which the move from A to A+ in the Mere Addition Paradox may be assessed according to them. The standard way of presenting this move assumes we depart from A and then we move to A+. But suppose we depart from A+, and are asked whether it would be worse to remain in A+ rather than to move to A, in which only those who are best off at A+ will exist. If we applied any of the three principles I have introduced above, we would then claim that remaining in A+ would be better. The reason would be that in A+ certain individuals whose existence is good exist, who instead would not exist if we moved to A.

But suppose we were not at A+, but at A. Many people have the intuition that if we are at A, while it would be fine if we moved to A+, it need not be necessarily better to move to A+ than to stay at A. The reason for this is that after all those extra individuals who would exist at A+ do not exist, in fact they have never existed and may never exist. Most people do not regret that there are possible happy individuals who may have lived but have never existed. On the contrary, if moving from A to A+ implied that the wellbeing of some of those in A is reduced, these principles would claim that A+ would be worse than A. Therefore, A might be better than A+ if A was actual or previ-

ous than $A+$, or if the identity of those in A was established but not the identity of those in $A+$. But $A+$ would always be better than A if $A+$ was actual or previous than A , or if the identity of those in $A+$ was established.

3.3. Why person-affective principles that seem to deny non-asymmetry fail to support an essentially comparative view

Given what we have just seen, it seems that these three principles I have just introduced appear to imply an Essentially Comparative View in an even more radical way than the Narrow Person-Affective View. This is so because although rejecting the transitivity of betterness is one way to reject the Essentially Comparative View, rejecting the asymmetry of betterness is another, more radical, way to do so. I will now claim, however, that this is not really the case. These person-affective principles only *seem* to imply that “better than” is asymmetric and intransitive. On closer examination, they are not.

To see this, we must consider that whether an outcome is better than another depends on which of those outcomes is the actual outcome. These facts need to be taken into account to know what outcome it would be better to obtain. There are two ways to explain this. We can count temporal position, actuality and the determination of individuals’ identity as something internal to outcomes, or we can count it as something external to them.

Suppose we consider that these circumstances are internal to outcomes themselves. That is, suppose they are part of that which defines a certain outcome and, therefore, of what can distinguish it from a different outcome. This means that two outcomes that are equal in everything except the time at which they occur are in fact different outcomes. And the same happens if one of them is actual and the other one is not, or if the identity of the individuals in them is determined beforehand or not. If we accept this we will have to conclude that A and B are not really the same outcomes when A is better than B and when “ B ” is better than “ A ”. And when we claim that A is better than B , B is better than C , and C is better than A , what happens is that A when A is better than B is different from “ A ” when C is better than “ A ”.

That is, let us assign the names A' and B' to “ A ” and “ B ” when “ B ” is prior to “ A ”, when “ B ” instead of “ A ” is actual, or when the identity of the individuals existing in “ B ”, but not in “ A ”, is determined. If $A > B$ and $B' > A'$, then $A \neq A'$ and $B \neq B'$.

Let us now assign the name B' to “ B ” when “ B ” is prior to C , when “ B ” is actual, or when the identity of the individuals existing in “ B ”, but not in “ C ”, is determined, and the names C' and A' to “ C ” when “ C ” is prior to “ A ”, when “ C ” is actual, or when the identity of the individuals existing in “ C ”, but not in “ A ”, is determined. If $A > B$, $B' > C$ and $C' > A'$ then $A \neq A'$, $B \neq B'$ and $C \neq C'$.

Given this, there is no reason to deny betterness is symmetric. But there is no reason either to deny it is a transitive relation.

Consider now the other solution. Let us suppose that when we compare two outcomes, whether one outcome is previous to another one, or is actual, or has individuals whose identity is already determined, is something “external”, so to speak, to whatever defines the outcomes themselves. We have seen that according to these principles it is impossible to know whether an outcome is better than another one without knowing which one is prior or actual, or whether the identity of those who exist in them is determined. But then, this means that knowing *everything* about two outcomes is not enough to know which one is better. This means that an outcome cannot be considered better or worse than another one *in itself*. So in order to compare two outcomes we need to know also something “external” to these outcomes.

If this is right, then whenever we compare which outcome is better according to the three principles I have presented we are not comparing the outcomes *as such*, but outcomes *and* something else. Therefore, it is not the outcomes *themselves* that are ordered according to the relation “all things considered better than”. We must thus conclude that these principles are compatible with (i) the view that “all things considered better than” is a transitive relation, and with (ii) an Internal Aspects View.

As we saw above, this is not the case when we consider other principles such as the Narrow Person-Affecting Principle whose application does not depend on factors such as time, actuality or whether the identity of individuals is fixed. Whenever we compare two outcomes according to principles such as this one we can tell which one is better without having to know anything else apart from the outcomes. Therefore, these principles never give rise to comparisons regarding betterness that are not asymmetric. But they can genuinely give rise to intransitive comparisons regarding betterness all things considered. Therefore, as long as we accept the Narrow Person-Affecting Principle (or any similar principle) we will have to reject the Internal Aspects View.

The difference between the Narrow Person-Affecting Principle and the three other principles I have presented is that, unlike these three principles, the Narrow Person-Affecting Principle compares different outcomes by virtue of the features of the outcomes themselves and that these features do not include circumstances, such as which outcome is previous, or actual. It is for this reason that according to the three other principles in certain circumstances A would be preferable to B and in other circumstances B would be preferable to A, while according to the Narrow Person-Affecting Principle this cannot be so.

This result, however, can also mean that while according to the Narrow Person-Affecting Principle it is possible that $A > B > C > A$ (though not that $A > B$ and $B > A$), this is not so according to the other three principles. While according to them it seems that it can be the case not only that $A > B > C > A$, but also that $A > B$ and $B > A$, this is so either because (a) when A is better than B, B and A are different than they are when “B” is better than “A”; and when A is better than B and B is better than C, A and C are different than they are when C is better than A; or (b) because the comparison is not only between outcomes but between outcomes *plus* other circumstances. It is due to this that the Time-Dependent Person-Affecting Principle, the Actuality-Dependent Person-Affecting Principle and the Identity-Dependent Person-Affecting Principle are fully compatible with an Internal Aspects View. This means it is possible to reject an impersonal view when it comes to comparing outcomes with different individuals and yet accept an Internal Aspects View. This setting aside the fact that those who accept an impersonal view will have no reason to reject the Internal Aspects View either.

4. WHAT HAPPENS IN SPECTRUM ARGUMENTS?

4.1. No essentially comparative principle applies in Spectrum Arguments

Let us assess another way Temkin defends an Essentially Comparative View, the one that appeals to Spectrum Arguments. We have seen already that in them our intuitions regarding which outcomes are better than others appear to imply that betterness is not transitive. However, this does not happen due to the application of a certain essentially comparative principle such as the Narrow Person-Affecting Principle. To be sure, in Spectrum Arguments one can assume an Essentially Comparative View. But if that is so, it needs to be due to reasons other than the application of essentially comparative arguments such as the ones presented in the previous sections (note that appeals to person-affecting reasons play no role here). In fact, what happens in Spectrum Arguments is that certain principles appear to clearly outweigh other principles in certain comparisons, but they appear to be clearly outweighed by them in other comparisons. To use the language Temkin introduces here (2012: ch. 2 and 5, appendices A and B), for some comparisons between outcomes there is a certain “standard view” that seems to be clearly right, while for other comparisons between outcomes there is another standard view that appears to be clearly correct as well.

Temkin points out that there is a standard view according to which trade-offs between the duration and the quality of a certain suffering are sometimes desirable (2012: 30). This is so because it is assumed that it is better to experience more intense suffering for a shorter period of time, rather than

less intense suffering for a longer period of time, if the difference in the intensity of the two pains is sufficiently small, and the difference in their durations is sufficiently large. This is the view we can think of when we face comparisons between immediate options, as that of A against B, B against C, C against D, etc. It is therefore the one that may drive us to think that $A > B > C$. When we compare options that are located far away in the spectrum, however, as when we compare Z against A, Temkin claims that there is another standard view according to which trade-offs between the duration and the quality of a certain suffering are not acceptable. This is so because it is assumed that it is worse to receive a more intense pain of some non trivial duration than a very mild pain of virtually any duration. This is the one that drive us to think $Z > A$.

So *two* standard views apply here, not a single principle.⁹ And note that none of these two views is an essentially comparative principle as defined above. Suppose we only accepted the standard view that claims that trade-offs between quality and duration can be desirable and therefore concludes that $A > B > C$. We would never reach any result according to which $A > B > C > \dots > Y > Z > A$, because that view would give us no reason to conclude that $Z > A$. Suppose now that we only accept the standard view that in some cases trade-offs between quality and duration are not desirable and concludes that $Z > A$. Again, the same result obtains: this view would never drive us to conclude $A > B > C > \dots > Y > Z > A$. In this case, because the view itself gives us no reason to accept $A > B > C$ or $Y > Z$. So, our intuitions may drive us to conclude that $A > B > C > \dots > Y > Z > A$ only when we combine these two principles.

Consider now the issue the other way around. Suppose we accept an essentially comparative principle such as the Narrow Person-Affecting Principle. This need not give us any reason to reject transitivity in Spectrum Arguments, because in them no variation regarding the individuals involved in different outcomes occurs. Therefore, if in these cases transitivity and the Internal Aspects View also fail to apply it must be due to completely differ-

9. Note that the contradictions between the application of principles sometimes works in different directions, so it is possible to draw two Spectrum Arguments in which things work just the other way around as in Temkin's main Spectrum Argument. For instance, consider one Spectrum Argument that starts with A, in which we suffer a very mild pain for a month. Then compare it to B, in which we suffer moderate pain for three weeks. B seems to be worse than A. Then consider C, which is a much worse pain for two weeks. C appears to be worse than B. And so on. Finally, you reach Z, where you suffer excruciating pain for a second. Many would deem Z preferable to A (though maybe not on reflection, providing a second is a relevant duration.).

The fact is that the intuitions we have towards each Spectrum Argument vary. For instance, I find the idea that excruciating pain for a second is worse than a headache for a month harder to accept than the idea that torture for a year is worse than some very mild pain lasting continuously for immense periods of time. But other people have different intuitions.

ent reasons. As I mentioned earlier, the only reasons to doubt about whether transitivity applies here is the conflict between our intuitions concerning the two different principles that we may accept in Spectrum Arguments. The examination of Spectrum Arguments must thus be different from the discussion of the cases involving essentially comparative principles visited above.

4.2. Why denying betterness is transitive does not solve the puzzles entailed by Spectrum Arguments

There are reasons to reject that the challenge that Spectrum Arguments present should drive us to reject the conclusion that transitivity does not apply to them. The main reason for this is that rejecting that betterness is transitive does not solve the problems implied by the conflict between different standard views.

Suppose we granted that in Spectrum Arguments comparisons must be essentially comparative, and that betterness can be intransitive. I may know that I intuitively prefer A to B, B to C, C to D... and then Y to Z and Z to A. And I can decide accordingly which outcomes are better when I compare two options in turn. But what happens if I do not only need to compare two options, but more? What happens, in particular, if I want to know what is the best option among all the available options? Rejecting transitivity may allow us to make the choices we find more intuitive when it comes to pairwise comparisons between different alternatives. But it offers no guidance whatsoever regarding what we may do when we have to choose *one* outcome in the whole spectrum. It leaves us without any way to look for a global maximum in which the harms we suffer are minimized (see Elster, 2000; McClennen, 1990: 231). If anything, it makes things far more complicated, since transitivity offers at least a possible method to solve the problem.

We may consider that this is not really crucial to the problem we face here. Having trouble finding the best outcome is surely a problem. But the fact that some outcomes such as Y and Z seem to be better than others such as A and B, which in turn seem to be better than others such as C and D, is also very problematic. So we could think that while the problem of which is the best outcome is one everyone faces, this other problem is one that only those who accept the Internal Aspects View have to face. In this way, rejecting transitivity would at least allow us to make some progress.

This, however, does not seem correct, because the problems we face when we have to choose the best option in the spectrum have the same origin as the problems we face when we compare alternatives that seem to be intransitive. The reason we have problems identifying the best option is that A appears to be better than B, B appears to be better than C and Z seems to

be better than A. It is all about the conflict between the prevalence of two different standard views. This is the reason why, in the spectrum, it seems that $Z > A$ while it also seems that $A > B > C > \dots > Y > Z$. It is also the reason why we are at a loss to identify the best option. In addition, we may take into account something Temkin considers when he wonders: “[w]ouldn’t the *best* alternative be the one that was best in comparison with *all* other *possible* alternatives, whether or not we might ever actually face them?” (2012: 470). This is a reasonable view, and if it is right, then in comparing two outcomes it makes perfect sense to compare them considering all the different options that there may be in a spectrum within which these two outcomes can be included.

We might believe that even if this is so, the problems we will have to face to do this will not be exactly the same ones that those faced by advocates of the Internal Aspects view. We may think that if we reject the claim that betterness is transitive we will have more alternatives available among which we can choose what solution is the best. But this is not so. To see this, let us examine what reasons we may have to support as sound candidates to the best outcome in the spectrum.¹⁰

- (a) *Duration prevails*. We may hold that the worst pain is always the one that lasts more. According to this A is the best option. This view will solve the whole spectrum problem by rejecting the applicability of the two standard views.
- (b) *Quality prevails*. We may also hold that the worst pain is always the most intense pain. If so, Z is the best option. Again, this would solve the spectrum problem by denying the applicability of the two standard views.
- (c) *Expected utility*. We may also hold that expected utility theory tells us how good or bad is each option in the spectrum. We would then choose A. Unlike in the case of the two other possible solutions we have just seen this criterion does not immediately dissolve the spectrum problem. On this view it makes sense to claim that there is intuitively less utility in A than in Z. But expected utility theory assumes that betterness is transitive. Due to this, it can revise this first judgement by acknowledging that it is in contradiction with agreeing that A is better than B, B better than C, and so on until we agree that Y is better than Z, and therefore conclude A is also better than Z.
- (d) *Critical level*. We may accept a capped model according to which there would be a certain critical level that ruled out as bad any outcome in which pain became too intense, if the critical level was set

10. See on this Handfield (forthcoming).

to rule out outcomes according to quality, or too persistent if the critical level was set to rule out outcomes according to duration. Accordingly, we could choose a point of the spectrum such as M, for example, at which the pain was not too intense or too long. We may think that rejecting transitivity makes it easier for us to accept this, because it would make it easier for us to accept any option different from A. But the fact is that we can accept any of those options even if we do not reject transitivity. Suppose we do this in the case of M. This will imply that we will have to accept that M is both better than L *and* N ($L < M > N$). This is very counter-intuitive and thus a high price to pay for holding this position. It is hard to see how M may be better than L and N, since the same reasons for M to be better than N seem to make it worse than L. But both advocates of transitivity and advocates of intransitivity face the same problem here.

- (e) *Strength of preference in pairwise comparisons.* We may also decide to choose the option for which our preference over the immediate next option in the spectrum is the strongest one. In that case we would choose Z, it seems, because our preference for Z over A is stronger than our preference for A over B, from B over C, etc. We may think that this solution is easy for advocates of intransitivity. But it is not. The reason is that if transitivity does not apply, then there is no reason to assume a certain ranking according to which a certain outcome comes after another one. So the whole idea of having an immediate next option in the spectrum ceases to make sense. But then, for advocates of intransitivity, the intuitive initial preference for Z over A is weaker than that for Y over A, and that of X over A. And, as I have just said, if transitivity does not apply, then there is no reason to assume that Z, and not Y or X, is the relevant outcome we must compare with A. And mind that both Y and X are clearly better than Z, yet similar enough to Z to make the distinction in quality between Z and A, and between Y to A trivial. So if we reject transitivity we will be at a loss about how to solve this problem.
- (f) *Special preferences.* We may just maintain an arbitrary view according to which some solution such as, say, G, is the best one just because it is. For instance, we may maintain that there's something special about mild torture for 10 years that makes it less bad both than a slightly worse torture for fewer years and a slightly milder torture for more years. Again, there is no reason why we may accept any solution if we reject transitivity but not if we accept it. As it happened when we considered critical levels, advocates of transitivity have to bite the bullet that some option is better than both the previous and the following option in the spectrum (which they may find preferable than having to reject transitivity). But advocates of

intransitivity need to do so as well if they are to give a solution to this problem at all.

There is only one way in which advocates of intransitivity can avoid all these difficulties which is not available to advocates of transitivity: by claiming that there is just no fact of the matter as to what outcome is the best one in the spectrum. But this is certainly not a way to solve the problem; rather, it is a way to claim (i) either that the problem has no solution or (ii) that we cannot solve it, both of which are odd replies.

In all these cases, *rejecting transitivity fails to facilitate a solution*. In fact, it makes it harder, if not impossible, to do so. To be sure, there can be other solutions to this problem apart from the ones I have presented. But it seems that the abovementioned problems both for advocates of transitivity and intransitivity persist.

5. WE DO NOT NEED TO REJECT THE INTERNAL ASPECTS VIEW

If my arguments above are correct, we have reasons to maintain transitivity and the Internal Aspects View. We have seen that some principles that appeal to person-affecting reasons are essentially comparative. But the intuitions supporting these principles can also be accommodated by means of other principles which are not essentially comparative. In addition, the argument will have no force for those who hold an impersonal view.

We have also seen that it is possible to try to explain our intuitions in Spectrum Arguments by rejecting transitivity. But this does not solve the problems these arguments pose, in light of which we may simply opt for not rejecting it and not accepting the Essentially Comparative View.

We can thus conclude that we do not need to reject the Internal Aspects View to find apt solutions to the challenges pressed against it. Moreover, rejecting transitivity fails to yield better solutions.

6. THE CONFLICT BETWEEN TEMKIN'S NORMATIVE AND METAETHICAL VIEWS

6.1. Can inconsistent moral intuitions track an objective reality?

The problems discussed thus far are normative and methodological, but may entail also metaethical consequences. This need not be so, since, strictly speaking, normative claims need not depend on metaethical views. Temkin's arguments, however, may give us reasons against moral realism (even though Temkin is a realist himself).

This is so because if there are some substantive moral claims that are true regardless of what we think, we have no epistemological access to the moral reality other than our intuitions. Moral realists thus claim that our intuitions somehow track moral reality. Temkin's attack on the Internal Aspects View, however, shows that the intuitions most of us have are inconsistent and chaotic. Could it be that moral reality is also inconsistent and chaotic? This seems implausible. Given this, it seems that either moral realism is also implausible or our moral intuitions do not really track any moral reality, in which case moral realism is indefensible (since we have nothing to back the claim that some moral reality exists).

Temkin argues that rejecting transitivity need not imply an inconsistent viewpoint (2012: 500). This, however, is besides the point. The issue is not the inconsistency involved in rejecting transitivity, or the Internal Aspects View. The issue is that Temkin's arguments show that most of us have inconsistent intuitions. To see this, consider the way Temkin argues that intransitivity does not entail inconsistency:

Suppose that there are three alternatives *A*, *B*, and *C*, such that we come to believe that it really *is* the case that, all things considered, *A* is better than *B*, *B* is better than *C*, but, all things considered, *A* is *not* better than *C*. Would this mean that we have inconsistent beliefs or that we thought the world was inconsistent?

No! It *would* mean this if we *also* thought that "all-things-considered better than" was a transitive relation.

Temkin's argument is not necessarily implausible. The problem, however is that many of us *do* think that "all-things-considered better than" is a transitive relation. I have presented several arguments defending this view. We can accept them and yet feel the strength of the intuitions Temkin appeals to when he claims, for instance, that in spectrum cases $A > B > C > D > \dots > Y > Z > A$. But that means we have inconsistent intuitions. This means that if our moral intuitions tracked the world, such world would also be inconsistent.

Temkin defends the compatibility of his normative position and realism, claiming that moral reality could be inconsistent. He writes that "[o]ur theories should reflect the world as it actually is, and on my view, whether or not the normative realm is vague, incomplete, or even inconsistent depends on facts about the normative realm, not on what is useful for us" (2012: 521). He adds to this: "[o]n my realistic conception of the normative realm, it is not [...] up to *us* to simply *decide* which positions should be accepted and which revised or rejected." (*ibid.*). Such statements leaves us without guidance to accept or reject any kind of normative or metaethical views. Moreover, we are the only ones who can decide which views to accept or reject, and we

have nothing but our intuitions and our capacity to compare them to make such decisions (this is the case of both realists and antirealists). Finally, the idea that there is an inconsistent moral reality appears to many of us both hardly conceivable and at any rate less likely than the falsehood of realism (regardless of other reasons we may have to doubt this).

6.2. Our intuitions towards metacriteria

How is it possible that there is this radical disagreement about the requirement of consistency between Temkin and (surely) many Temkin readers like myself? To answer this question we can examine the kind of intuitions with which we can appraise the apparent paradoxes that Temkin presents.

It seems that most of us have intuitions both about whether certain particular outcomes are better than others and towards certain general axiological principles. Similarly, we have intuitions about what we should do in some particular cases and about general normative principles. Some of us, however, have *also* intuitions about metacriteria concerning our axiological and normative theories. Some of us have a very strong intuition that betterness must be transitive. We also have a very strong intuition that our axiological and our moral views must be consistent. These are not the same; and it is possible that rejecting transitivity does not entail any kind of inconsistency. But they are both strong intuitive metacriteria for many of us. These are not, moreover, the only intuitions about metacriteria we have: many of us also have the intuition that there is something wrong with a theory that is very complex and includes lots of exemptions and provisos. In fact, many of us find these intuitions so compelling that when we have to decide between abandoning them or abandoning the claim that a year of torture must be worse than a very large number of mosquito bites we give up the claim about mosquito bites. However, Temkin and others find that decision unacceptable. This is because their intuitions regarding consistency, transitivity and simplicity are weaker or even nonexistent. Temkin is not only willing to give up transitivity and simplicity, but also consistency if it clash with his intuitions about pairwise comparisons of outcomes (2012: section 14.6). He argues that “[n]othing can *force* someone to give up a set of inconsistent views” (2012: 520).

Because we have inconsistent views, Temkin suggests we give up our aspirations to have a consistent approach, so we can keep all our views. This can help Temkin not only to hold the normative views he finds intuitive, but also to do so without abandoning moral realism. If we strongly believe all our views must be consistent, however, the problem continues, and consistency remains a requirement we do not want to drop. Moreover, we probably do not want to drop other metacriteria such as transitivity and simplicity. This gives us reasons to resist Temkin’s attack on the Internal Aspects View, and, if

the argument presented above is correct, it also present a serious challenge to his realist views in metaethics.

BIBLIOGRAPHY

- Binmore, K. G., and Voorhoeve, A., 2003: "Defending Transitivity Against Zeno's Paradox", *Philosophy and Public Affairs*, 31: 272-279.
- Broome, J., 2004: *Weighing Lives*, Oxford: Oxford University Press.
- Dancy, J., 2005: "Essentially Comparative Concepts", *Journal of Ethics and Social Philosophy*, 1, URL = <http://www.jesp.org/PDF/Dancy.pdf>.
- Elster, J., 2000: *Ulysses Unbound: Studies in Rationality, Precommitment, and Constraints*, Cambridge: Cambridge University Press.
- Friedman, A., 2009: "Intransitive Ethics", *Journal of Moral Philosophy*, 6: 277-297.
- Handfield, T. forthcoming: "Rational Choice and the Transitivity of Betterness", *Philosophy and Phenomenological Research*.
- Huemer, M., 2013: "Comparative Value, and the Methods of Ethics", *Ethics*, 123: 318-345.
- Norcross, A., 1997: "Comparing Harms: Headaches and Human Lives", *Philosophy and Public Affairs*, 26: 135-167.
- 1998: "Great Harms From Small Benefits Grow: How Death Can Be Outweighed by Headaches", *Analysis*, 58: 152-158.
- 1999: "Intransitivity and the Person-Affecting Principle", *Philosophy and Phenomenological Research*, 59: 769-776.
- McClennen, E. F., 1990: *Rationality and dynamic choice: Foundational explorations*. Cambridge University Press.
- Quinn, W., 1990: "The Puzzle of the Self-Torturer", *Philosophical Studies*, 59: 79-90.
- Rachels, S., 1998: "Counterexamples to the Transitivity of Better Than", *Australasian Journal of Philosophy*, 76: 71-83.
- Rachels, S., 2001: "A Set of Solutions to Parfit's Problems", *Noûs*, 35: 214-238.
- 2004: "Repugnance or Intransitivity: A Repugnant but Forced Choice", in *The Repugnant Conclusion: Essays on Population Ethics*, ed. J. Ryberg and T. Tännsjö, 163-186, Dordrecht: Kluwer.
- Temkin, L. S., 1987: "Intransitivity and the Mere Addition Paradox", *Philosophy & Public Affairs*, 16: 138-187.
- Temkin, L. S., 1996: "A Continuum Argument for Intransitivity", *Philosophy & Public Affairs*, 25: 175-210.
- 1997: "Rethinking the Good, Moral Ideals and the Nature of Practical Reasoning", in *Reading Parfit*, ed. J. Dancy, 290-234, Oxford, Blackwell.
- 2010: "Intransitivity and the Person-Affecting Principle", *Philosophy and Phenomenological Research*, 59: 777-784.
- 2012: *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*, Oxford: Oxford University Press.
- Voorhoeve, A., 2008: "Heuristics and Biases in a Purported Counter-Example to the Acyclicity of 'Better Than'", *Politics, Philosophy and Economics*, 7: 285-299.
- Weber, M., 2007: "Is Equality Essentially Comparative?", *Ethical Theory and Moral Practice*, 10: 209-226.

On the Asymmetry of “Better than”, Person-Affecting Views, Spectrum Arguments, the “Unhelpfulness” of Rejecting Transitivity, and Implications for Moral Realism

LARRY TEMKIN

Rutgers University

Abstract

This article responds to Oscar Horta’s article “In Defense of the Internal Aspects View: Person-Affecting Reasons, Spectrum Cases and Inconsistent Intuitions”. I begin by noting various points of agreement with Horta. I agree that the “better than relation” is asymmetric, and point out that this will be so on an Essentially Comparative View as well as on an Internal Aspects View. I also agree that there are various possible *Person-Affecting Principles*, other than the one my book focuses on, that people might find plausible, and that in some circumstances, at least, these might have deontological, rather than axiological significance. In particular, I grant that Horta’s *Actuality-Dependent Person-Affecting Principle*, his *Time-Dependent Person-Affecting Principle*, and his *Identity-Dependent Person-Affecting Principle*, might each be relevant to what we ought to *do*, without necessarily being relevant to which of two outcomes is *better*. But I reject Horta’s claim that essentially comparative principles don’t apply in Spectrum Arguments. I also argue against Horta’s view that the two Standard Views that underlie our intuitions in Spectrum Arguments are contradictory. I question Horta’s (seeming) position that there is no point in rejecting the transitivity of the “better than” relation on the basis of Spectrum Arguments, on the grounds that doing so won’t *solve* the predicament that Spectrum Arguments pose. Finally, I conclude my paper by challenging Horta’s interesting contention that my views about nontransitivity support an anti-realist metaethics, and are incompatible with the sort of realist approach to metaethics that I favor.

Keywords: Transitivity, Spectrum Arguments, Person-Affecting Principles, Internal Aspects View, Better than, Essentially Comparative View, Realism, Anti-realism, Sophie’s Choice, Moral Dilemmas.

Let me begin by acknowledging my gratitude to Oscar Horta for his thoughtful and sensitive comments in his article “In Defense of the Internal Aspects View: Person-Affecting Reasons, Spectrum Cases and Inconsistent Intuitions” (Horta 2014), and also for the generous spirit he displayed in presenting them. I will divide my responses into two main parts. In part I, I will note some points where I agree with Horta. In part II, I will note some points about which we disagree.¹

1.

In section 3, Horta discusses different possible types of Person-Affecting Principles, and the question of whether better than can be non-asymmetric. As Horta recognizes, the points he makes in this section are not in tension with my book’s claims. But I agree that they offer useful lessons to bear in mind as we try to determine what needs to be said about the “better than” relation and other analogous relations.

One of Horta’s main claims in section 3 is that the “better than” relation is asymmetric: so if, in any given context, A is better than B, *all things considered*, then it *can’t* also be the case that, in that very same context, B is better than A, all things considered. Even I, who am open to rejecting the Axiom of Transitivity for the “better than” relation, don’t reject the fact that the better than relation is asymmetric! Similarly, while I am open to rejecting the Axiom of Transitivity for the “equally as good as” relation, I accept the standard view that the “equally as good as” relation is symmetric: so if, in any given context, A is equally as good as B, *all things considered*, then it must also be the case that, in that very same context, B is equally as good as A, all things considered.

1. This article was originally written in response to the talk that Horta presented at the LEAP symposium on my book, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning* (Temkin, 2012), at Pompeu Fabra University in Fall 2012, and to an early draft of his article based on that talk. Unfortunately, shortly before this journal was to go to press, I received the final, revised, version of Horta’s article, and I was somewhat surprised to see that he had substantially revised his article, both in terms of adding new material that I hadn’t previously seen, and deemphasizing, or removing, some key claims or passages to which my original article had objected. I have, where possible, adjusted my article in light of Horta’s final changes. However, given the journal’s time constraints, there were some important aspects of the final version of Horta’s article to which I was unable to respond. In addition, in some cases I have thought it worthwhile to retain points that I raised with respect to his original talk and draft, since others may be attracted to views similar to those he previously held, even if Horta, himself, has now changed his mind on the matters. Thus, I acknowledge, here, that some of the points that I will be making are less relevant, or even not relevant, to his published article, and that, in some cases, when I highlight a supposed disagreement between us, the disagreement between us may have lessened, or disappeared altogether, since Horta first presented his views on these topics. To aid the reader, I try to make it plain in the text where I am mainly responding to his talk, as opposed to his published article.

To these uncontroversial claims, I would simply point out that the explanation I give for *why* the various Axioms of Transitivity could fail to hold, does not similarly challenge these other standard claims. In challenging the Axioms of Transitivity, I noted that some of the ideals people most value are *Essentially Comparative*. On an *Essentially Comparative View*, the factors that are relevant and significant for assessing an outcome may vary depending on the alternative with which it is compared. This opens up the possibility that, in any given context, the factors that are relevant and significant for comparing A with B, or B with C, may be different from the factors that are relevant and significant for comparing A with C. From this it follows that, in a given context, A might be better than (or equally as good as) B, in terms of *all* of the factors that are relevant and significant for making *that* comparison, and B might be better than (or equally as good as) C, in terms of *all* of the factors that are relevant and significant for making *that* comparison, and yet A might *not* be better than (or equally as good as) C, in terms of *all* of the factors that are relevant and significant for making *that* comparison. Thus, I claimed that on an Essentially Comparative View of ideals, the Axioms of Transitivity regarding the “better than” and “equally as good as” relations may fail, or fail to apply across different sets of alternatives to which we might have thought they should apply.²

However, as indicated, the explanation I give for *why* the Axioms of Transitivity may fail, or fail to apply, offers no reason to doubt the asymmetry of the “better than” relation, or the symmetry of the “equally as good as” relation. This is because whether one accepts an *Internal Aspects View* (according to which how good an outcome is depends solely on the internal features of that outcome), or an Essentially Comparative View of ideals, if, in a given context, A is better than B in terms of *all* of the factors that are relevant and significant for comparing A and B in that context, then it *will* be the case that B is *worse* than (and hence *not* better than!) A in terms *all* of the factors that are relevant and significant for comparing A and B in that context. Thus, “better than” is asymmetric, precisely as Horta has claimed and as I, and everyone else, should readily accept.

2. Some people believe that on an Essentially Comparative View the Axioms of Transitivity fail to hold, so that they should be rejected. Others insist that there are various ways of preserving the Axiom of Transitivity even on an Essentially Comparative View, so that they never fail, but that the Axioms of Transitivity may *fail to apply* in those cases where Essentially Comparative ideals are relevant for assessing different alternatives. In my book, I use the notion of *non-transitivity* to cover both the cases where we think the Axioms of Transitivity fail, and those where we think they fail to apply across different sets of alternatives to which we might have expected that they should apply. I argue that there are significant practical and theoretical implications of the “all-things-considered better than”, “equally as good as”, and “at least as good as” relations being non-transitive, whether or not this is because the relations fail, or “merely” because they fail to apply in the cases I discuss (Temkin, 2014: [currently, note 19, page number ?]). See also Temkin, 2012: 5, 16-8, 59-60, 66, 163-182, 197-8, 203-214, 223-5, and ch. 13.

Similarly, whether one accepts an Internal Aspects View or an Essentially Comparative View of ideals, if, in a given context, A is equally as good as B in terms of *all* of the factors that are relevant and significant for comparing A and B in that context, then it *will* be the case that B is equally as good as A in terms *all* of the factors that are relevant and significant for comparing A and B in that context. Hence, the “equally as good as” relation is symmetric.

In sum, I have offered an account of why the Axioms of Transitivity might fail to hold, but that account does nothing to challenge the uncontroversial claims that “better than” and “equally as good as” are asymmetric and symmetric relations, respectively.

Let me turn next to Horta’s discussion of different Person-Affecting Principles. Horta distinguishes between three different kinds of Person-Affecting Principles:

The Actuality-Dependent Person-Affecting Principle: In assessing possible outcomes, one should focus on the status of those who exist in the actual world, with the aim of wanting them to be as well off as possible, and (2) ignore the status of those who do not exist in the actual world, except that one wants to avoid harming them as much as possible.

The Time-Dependent Person-Affecting Principle: In assessing possible outcomes that haven’t occurred yet, one should focus on those who will exist in the outcome that will occur first, with the aim of wanting them to be as well off as possible, and (2) ignore the status of those who will not exist in the outcome that will occur first, except that one wants to avoid harming them as much as possible.

The Identity-Dependent Person-Affecting Principle: In assessing possible outcomes, one should focus on the status of those whose identity is already determined, with the aim of wanting them to be as well off as possible, and (2) ignore the status of those whose identity is not determined yet, except that one wants to avoid harming them as much as possible (Horta 2014: 10-11 of his draft).

I should mention that none of Horta’s versions of Person-Affecting Principles correspond exactly to the two versions I discuss in my book, my *Narrow Person-Affecting View* and my *Wide Person-Affecting View* (Temkin, 2012: 416-45), but this need not concern us here.

In his original talk, Horta contended that while positions like the Actuality-Dependent, Time-Dependent, and Identity-Dependent Person-Affecting Principles might be relevant, in certain circumstances, to our assessment of what we ought or ought not to *do* in choosing between two alternatives, it doesn’t follow that such judgments are tracking which of the two alterna-

tives, considered just by *themselves*, is *better, per se*. Here, too, I agree with Horta.

Consider, for example, Diagram One.

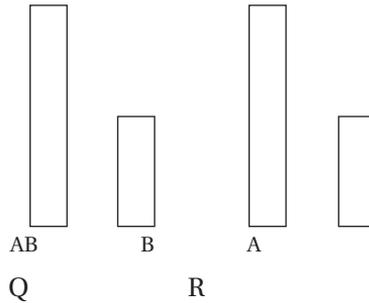


Diagram one

As drawn, Diagram One represents two possible outcomes, Q and R. Each outcome contains two groups, a better-off group and a worse-off group. There is no difference between the levels or number of people in the two better off groups, and similarly, for the two worse-off groups. The only difference concerns the *identities* of the better- and worse-off groups in the two outcomes. In Q, the A people occupy the better-off group and the B people occupy the worse-off group; in R, the reverse is true.

Assuming there are no morally relevant differences between the A and B people—so, for example, each person is equally talented, hardworking, deserving, and so on—it seems clear that, considered just by *themselves*, Q and R are *equally good*. But now, suppose that there were a presently existing *actual* outcome, with people whose identities were already determined, and it looked like P in Diagram Two.

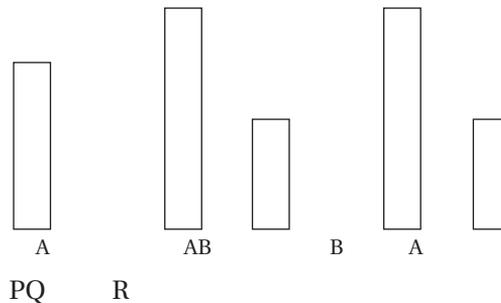


Diagram two

In P, a presently existing actual outcome, the A people already exist, and they are at a level between that of the better- and worse-off groups in possible future outcomes Q and R. Suppose, next, that we could transform the P outcome into one like Q or R. That is, we could either raise the A people up, and bring another group of people, B, into existence at a lower level, so as to produce an outcome like Q, or, alternatively, we could bring another group of people, B, into existence at a high level, but at the cost of lowering the A people, so as to produce an outcome like R.

In accordance with Horta's Actuality-Dependent, Time-Dependent, and Identity-Dependent Person-Affecting Principles, it would be permissible, desirable, and perhaps even *obligatory* to bring about Q, and impermissible, undesirable, and prohibited to bring about R. But even if this showed that P would be *improved* by being changed into Q and *worsened* by being changed into R, that *doesn't* entail that Q, considered just by *itself*, is a worse outcome than R, considered just by *itself*. This should be evident, if one considers the fact that instead of the third alternative P, being a presently existing actual outcome, it *might* have been a fourth alternative, O, that was a presently existing actual outcome, where O was *just like* P except that instead of the A people existing at a level between the better—and worse—off groups in Q and R, the B people existed at that level. In that case, it would have been true that on all three of Horta's Dependent Person-Affecting Principles, O would be *improved* by being changed into R, and it would be *worsened* by being changed into Q. Thus, we would have to abandon the non-asymmetry of the "better than" relation *if* we thought that we *could* infer that Q was in *itself* better than R, simply from the fact that it *would* be desirable to transform P into Q, but undesirable to transform P into R; because the same reasoning would then entail that R was in *itself* better than Q, since it would *also* be desirable to transform O into R, but undesirable to transform O into Q.

A fortiori, as Horta contended in his talk, even if there are cases where in accordance with any of his Dependent Person-Affecting Principles it would be obligatory to bring about one outcome, O_1 , rather than another outcome, O_2 , it doesn't follow that the judgments yielded by such principles support the conclusion that, considered just by *themselves*, O_1 is better than O_2 . However, let me add that Horta's position is not merely compatible with my claims in *Rethinking the Good*, it follows directly from the fact that Horta's three Person-Affecting Views are Essentially Comparative as I characterized that notion. As noted previously, on an Essentially Comparative View, different factors can be relevant and significant for assessing the relative goodness of outcomes like Q or R, depending on the alternatives with which they are compared. Thus, Q might have one value in comparison with R when those are the only alternatives, but a different value in comparison with R

when each is itself an alternative to some third outcome P, and a still different value in comparison with R when each is itself an alternative to some fourth outcome O.

2.

Let me turn next to some points of disagreement with Horta, some of which are fairly minor, but others of which are not.

First, in discussing my Spectrum Arguments, Horta titles his subsection 4.1 “*No essentially comparative principle applies in Spectrum Arguments*”, and he calls special attention to the claim “that appeals to person-affecting reasons play no role here” (Horta, 2014: 14). I agree with Horta’s claim that Person-Affecting Views do not underlie my Spectrum Arguments.³ However, I take exception to the title of his subsection, which, I believe, is either misleading or mistaken.

As I have presented and analyzed them, Spectrum Arguments *do* arise because of the *Essentially Comparative View* of ideals. Consider, for example, my Spectrum Argument where the first member of the Spectrum involves a very long life with 15 mosquito bites per month and two years of torture, and the last member involves a very long life with 16 mosquito bites per month but no torture. I claimed that two distinct views guided our thinking in making different comparisons along the Spectrum. The *First Standard View* reflects an *Additive Aggregationist Approach* and the *Second Standard View* reflects an *Anti-Additive-Aggregationist Approach*. Both Views are limited in scope, in that they seem relevant and significant for making certain comparisons but not others. In particular, I pointed out that where the differences in the intensity of pains between two alternatives was very small, the First Standard View seemed plausible and appropriate for comparing those alternatives, so, in particular, it seemed appropriate for comparing my Spectrum’s first alternative with the second, the second with the third, the third with the fourth, and so on. On the other hand, where the differences in the intensity of pains between two alternatives was very large, the Second Standard View seemed plausible and appropriate for comparing *those* alternatives, so, in particular, it seemed appropriate for comparing my Spectrum’s first few alternatives with its last few alternatives. Thus, the factors that seemed relevant and significant for assessing the Spectrum’s first alternative were *different* depending on whether it was being compared with the Spectrum’s second alternative or its last alternative. This is in keeping with the *Essentially Comparative View* of ideals, and is at odds with the *Internal*

3. However, like the Spectrum Arguments, Person-Affecting Views challenge the Axioms of Transitivity. See Temkin, 2014: sec. V; and Temkin, 2012: ch. 12.

Aspects View of ideals (see Temkin, 2014: section III, pp. 18-19 of draft; and Temkin, 2012: 62-6, 229-31, 369-74).

In sum, while it may be true that there is no *single* Essentially Comparative ideal like a Person-Affecting View that underlies my Spectrum Arguments, I think it *is* true that an Essentially Comparative approach best explains what is going on in my Spectrum Arguments, and why they ultimately put pressure on the Axioms of Transitivity.

Let me turn to a second point. Horta suggests that my two Standard Views are *contradictory*.⁴ I deny this.

The First and Second Standard Views would be *contradictory* if there were any particular judgments which the First Standard View made which were denied by the Second, or vice versa. But this is not, I think, the case. For example, where the First Standard View yields the judgment that the first member of my Spectrum is better than the second, the Second Standard View doesn't *deny* this judgment, rather it is silent. Specifically, as I characterized it in my book, the Second Standard View simply doesn't *apply* for comparisons involving such alternatives. Likewise, where the Second Standard View yields the judgment that the Spectrum's first member is worse than the last, the First Standard View doesn't *deny* this judgment, rather it is silent. Again, as I characterized it, the First Standard View simply doesn't *apply* for comparisons involving such alternatives.

Note, there would be nothing *contradictory* about the claims that John is in love with Mary and that John isn't in love with Tim, *even if* Mary is in love with Tim. There would also be nothing problematic about such claims, and this for the simple reason that "is in love with" isn't a transitive relation! Likewise, there will be nothing *contradictory* or even problematic about the claim that Spectrum's first outcome is better than the second, but the first is *not* better than the last, *even if*, for each pair of adjacent outcomes n and $n + 1$ along the Spectrum, n is better than $n + 1$, *as long as* "all-things-considered better than" *isn't* a transitive relation.

But, of course, *if*, as many believe, the First Standard View *is* relevant and significant for comparing adjacent outcomes along my Spectrum, but doesn't apply for comparing the first and last outcomes, and *if*, as many also believe, the Second Standard View *is* relevant and significant for comparing my Spectrum's first and last outcomes, but doesn't apply for comparing

4. Horta was clearer about this in his talk, than he is in his article, where he seems to de-emphasize this claim. However, he still seems to hold the view in question. See, for example, note 10 of Horta, 2014: [p. 15 of draft, where he writes "Note that the *contradictions* between the applications of the [Spectrum Arguments'] principles sometimes works in different directions... (emphasis added)".

my Spectrum’s adjacent outcomes, then there is good reason to believe that “all-things-considered better than” *isn’t* a transitive relation, in which case the First and Second Standard Views won’t be contradictory, and neither will their respective judgments that the Spectrum’s first outcome is better than the second, but not better than the last.

Third, Horta suggests that there is a powerful reason to resist being driven to “reject the conclusion that transitivity does not apply to [Spectrum cases, since]... rejecting that betterness is transitive *does not solve the problems* implied by the conflict between different standard views (emphasis added)” (Horta, 2014: 16 of final draft). Later, Horta emphasizes that “*rejecting transitivity fails to facilitate a solution*. In fact, it make it harder, if not impossible, to do so” (Horta, 2014: 19 of final draft). In essence, then, Horta believes that there is no *point* in rejecting transitivity on the basis of Spectrum Arguments, if doing so won’t help us to *solve* such arguments.⁵

To a large extent, I agree with these remarks. Indeed, I emphasize some of these very same sentiments in my book, when I’m discussing the costs and benefits of accepting or rejecting the different positions underlying my impossibility arguments. But, I’m not sure, exactly, what is supposed to *follow* from such observations.

I am a philosopher seeking the truth. We would very much *like* to be able to answer certain questions in a certain way. But what if the sad truth is that the answers we seek are not to be found, or cannot be answered in the way we had thought or hoped. Is it not an important advance in our understanding of the normative realm if we learn that this is so?

I am reminded here of the Socratic claim regarding why the Oracle called him the wisest of all men. It was, Socrates claimed, because whereas most people thought that they knew a lot, and were wrong, he knew that he knew nothing (other than the fact that he knew nothing!).⁶ As Socrates recognized, it can be as important to know what we don’t know, and to learn what we can’t know, as to continue to seek solutions along a path where they can’t be found.

This raises a related point. Many years ago, when I first began thinking about and teaching these issues, Carl Hoefer, who was then still an undergraduate, was quite pleased by my results, and the implications he took them to have. Hoefer was worried about the dominance of consequentialist reasoning in much of contemporary normative reasoning. At the time,

5. Horta was, I think, even clearer about this in his talk, where he really emphasized this point.

6. The position in question is often attributed to Socrates on the basis of a key passage (21d) of Plato’s *Apology*.

Hoefler thought it would be better if moral philosophers spent more time focusing on considerations of character, of the sort championed by Aristotle, or on deontological considerations, of the sort championed by Kant. For Hoefler, my results suggested a vindication of sorts for those who thought that a focus on consequences, and in particular on the aim of bringing about the best available outcome by maximizing the good, was the wrong way to proceed in determining how we ought to act, morally.

Now I don't abandon the appeal to consequences lightly, if at all. Indeed, as I emphasize in my book, I don't really even see how that could be a possible option, and I am certain that there would be *enormous* costs to such a move. Moreover, and more troubling, I'm not sure that virtue-based or deontological-based reasoning can wholly avoid the sorts of worries that arise in my book from Spectrum Arguments or an Essentially Comparative View of ideals. But having said all that, the mere fact that abandoning the transitivity of the "all-things-considered better than" relation wouldn't help us to decide what to do if, for example, we found ourselves facing a Spectrum-type choice and we wanted to bring about the best available outcome, *doesn't* show us that transitivity *shouldn't* be abandoned. Perhaps, reluctantly, it should. However intuitively unpalatable, this is an alternative that requires careful consideration. Indeed, perhaps taking such an option seriously will force us to pursue other paths in our exploration of the normative realm that may ultimately prove to be more fruitful than the paths on which most moral philosophers have focused up until now.

Let me conclude my response to Horta with some comments on the topic of section 6 of his article, moral realism.

As I note in my book, Derek Parfit once claimed that *if* my arguments were sound, they amounted to the most skeptical argument against moral realism since David Hume's arguments.⁷ Since neither Parfit nor I are moral skeptics, that gave us both reasons to *hope* that my arguments aren't sound—at least, if Parfit were right in his assessment about their implications. But, of course, hoping doesn't make it so! Moreover, as my previous comment suggests, how devastating my results may prove to be will ultimately turn on whether other fruitful paths in the normative realm might be found that do not fall victim to my arguments.

A key question to be addressed concerns how much of the normative realm depends on our being able to provide a coherent ordering, in the form of a transitive ranking, of outcomes or choices. Even if, in the end, there is no meaningful transitive ranking of outcomes that we can correctly appeal to in

7. Actually, Parfit has made this claim to me on multiple occasions over the years during discussions about my work.

our normative deliberations, it is arguable that there may still be many full-blooded realist considerations that would rightly have a bearing on what choices we ought, morally, to make in the living of our lives. Perhaps some of these would be deontic-, caring-, or virtue-based in nature. But, as implied above, perhaps some of these would be along new lines yet to be discovered and developed; lines which, perhaps, will only be discovered if we are forced to look in new directions for navigating the normative realm.

Finally, I am a realist in thinking that *if* we end up abandoning the Axioms of Transitivity, it will be because we are recognizing that there are compelling reasons to accept Essentially Comparative principles like the Narrow Person-Affecting View, the Pareto Principle, person-affecting versions of Utility and Maximin, and combinations of positions like the First and Second Standard Views. That is, I believe that there may be good *reasons* to abandon the Axioms of Transitivity, and that we will only be driven to such a position by the force of such reasons. So, this makes me a *realist* about reasons, as opposed to a skeptical anti-realist who denies that there could be *reasons* to accept or reject *any* particular principles, including principles of consistency like the Axioms of Transitivity.

My position here is similar to that of some people who believe in the possibility of genuine moral dilemmas. Consider one classic example of a so-called moral dilemma, *Sophie's Choice* (Styron, 1979). Sophie seemingly faced three choices: she could save her son, in which case the Nazis would murder her daughter, she could save her daughter, in which case the Nazis would murder her son, or she could do nothing, in which case the Nazis would murder both her children.

Anti-realists about reasons believe that in the most fundamental sense it doesn't matter what Sophie chooses. Specifically, they believe that, ultimately, there are *no reasons* of any kind guiding Sophie's choice, or any other choice for that matter. On the anti-realist position, there is no reason either way for Sophie to save her son, save her daughter, allow both to be killed, or, for that matter, to pursue a fourth option of killing both of her children herself, and perhaps a lot of other innocent victims who would otherwise have survived!

Those realists about reason who believe in the possibility of genuine moral dilemmas (many realists do not), believe that there are overwhelmingly compelling reasons for Sophie to save her son, but that there are *also* overwhelmingly compelling reasons for Sophie to save her daughter, and that the nature and structure of the reasons in question are such that they do not, and cannot, balance or cancel each other out. On their view, the fact that she can't save *both* of her children doesn't alter the fact that Sophie *should* save her son, which she *can* do, *and* that she *should* save her daughter, which she

also can do. Accordingly, for such moral realists, Sophie is facing a moral blind alley, or moral dilemma, in the sense that whether she chooses to save her son or her daughter she will, in a deep and fundamental way, have acted wrongly. On this view, whatever choice Sophie makes, she will have unavoidably acted contrary to compelling reasons for acting otherwise than she did, reasons which were not cancelled out or balanced by the likewise compelling reasons on which she chose to act.

Importantly, such realists would vehemently deny that there are *no* reasons applicable to Sophie's choice. Contrary to the anti-realist, they would insist that Sophie has strong reason not to let *both* children be killed, and even stronger reason not to kill her children herself along with a number of other innocent victims who would otherwise survive. But they believe that the very real and compelling reasons that have a bearing on her situation put Sophie in a "no win" situation. She *must* choose to save her son or her daughter, but whatever choice she makes will be wrong! On the view in question, moral dilemmas are truly tragic situations from which there is no rational escape. But they only obtain because there genuinely *are* moral reasons whose nature and structure give rise to such dilemmas.

My point in discussing moral dilemmas is not to defend the view that there are such dilemmas, but to illustrate how such a view is consistent with a realist view of reasons. Similar thinking applies, I believe, regarding my Spectrum Arguments, or the other arguments I have given threatening the Axioms of Transitivity. One may be a realist about reasons, but believe that an Essentially Comparative View of ideals is true. *If* the Essentially Comparative View is true, then there may be no transitive ordering of the alternatives in my Spectrum Arguments, or of many of the other sets of alternatives my book discusses. But this won't be because the anti-realists are right. Rather, it will be because of the nature and structure of the genuine reasons that exist, and that bear on the alternatives in question.

Or so I believe, anyway.

BIBLIOGRAPHY

- Horta, O., 2014: "In Defense of the Internal Aspects View: Person-Affecting Reasons, Spectrum Cases and Inconsistent Intuitions", *Law, Ethics and Philosophy* 2: ????.
- Styron, W., 1979: *Sophie's Choice*, New York: Random House.
- Temkin, L., 2012: *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*, Oxford: Oxford University Press.
- 2014: "Rethinking the Good - A Small Taste", *Law, Ethics and Philosophy* 2: ????.

Internal or External Grounds for the Nontransitivity of “Better/Worse than”

INGMAR PERSSON

Gothenburg University and Oxford Uehiro Center for Practical Ethics

Abstract

In his book *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning* Larry Temkin contrasts two views of ideals for evaluating outcomes: the Internal Aspects View and the Essentially Comparative View. He claims that the latter view can make the relation of being better/worse than all-things considered nontransitive, while the former can't. This paper argues that the Internal Aspects View can also be a source of nontransitivity. The gist of the argument is that perfect similarity as regards supervenient properties, like value, is compatible with differences as regards their subvenient properties and that it's logically possible that such sets of insufficient differences add up to differences that are sufficient for supervenient differences. Thus, perfect similarity or identity is nontransitive as regards the supervenient property of value, and this implies that the relation of being better/worse than all things considered is also nontransitive.

Keywords: Derek Parfit, Larry Temkin, Transitivity, Non-transitivity, Supervenience, Outcome value, Identity, Similarity.

1. INTERNAL AND COMPARATIVE VIEWS AND THE NONTRANSITIVITY OF “BETTER THAN”

Larry Temkin's monumental book *Rethinking the Good* is by far the most resourceful and penetrating investigation into the various aspects of the value of outcomes to date. It's therefore invaluable to anyone with an interest in these matters. A central theme is the contrast between two views of ideals for evaluating outcomes: the Internal Aspects View and the Essentially Comparative View. According to the Internal View, how good an outcome is with respect to any relevant ideal depends solely on its *internal* features, that is, features that it has irrespective of its relations to other outcomes (which are not parts of it). Consequently, how good an outcome is all things —i.e. all relevant ideals— considered, will depend entirely on its internal features. According to the Comparative View, there are some ideals such that how good

an outcome is with respect to these ideals depends not only on its internal features, but also on its relations to other outcomes with which it is compared. Thus, how good an outcome is all things considered will depend also on its external relations.

To illustrate, consider three outcomes:

- A: a large number of people at a high level of welfare;
- A+: this population plus an equally large number of individuals at a significantly lower level of well-being, but still well above the neutral or zero level; and
- B: these two populations at a level which is a bit higher than halfway between their levels in A+.

Now suppose that inequality isn't a bad feature that detracts from the value of an outcome if it comes about by bringing individuals into existence. Hence, it isn't a respect in which A+ is worse than A. However, it is a respect in which A+ is worse than B, since these outcomes only contain individuals who exist in both outcomes. Therefore, B would be better than A+ with respect to equality, but A+ would not be worse than A in this respect. So, if A+ is better than A because its sum of welfare is greater, A+ will plausibly be better than A all things considered. B will reasonably be better than A+ all things considered because it is better both in terms of equality and in terms of aggregate welfare. True, it's worse in one respect because nobody in B is as well off as the better-off individuals in A+; so, some of the better things in life may be lost. But we might feel that this aspect is outweighed by the other two aspects in which B is better. Nonetheless, we might also feel that B isn't better all things considered than A because its greater sum of well-being doesn't outweigh the qualitative loss. Then we would face what Derek Parfit calls *the Mere Addition Paradox* (1984: Ch. 19) if we also endorse the transitivity of the relation of being better than all things considered which implies that B is better all things considered than A because B is better all-things-considered than A+ which is better all things considered than A.

By adopting a Comparative View of the ideal of equality and taking it to be relevant for the comparison between A+ and B, but not for the comparison between A and A+, we can remove the paradox by denying the transitivity of "all things considered better than". Accordingly, Temkin associates the nontransitivity of this relation with the Comparative View. In contrast he affirms that on the Internal View "all things considered better than" *must* be a transitive relation" (2012: 494).¹ I shall argue, however, that a rejection of transitivity can also be justified by the Internal View. The source of it must

1. However, in a footnote he seemingly concedes the possibility of constructing an Internal View that doesn't imply transitivity, but he doesn't explore this possibility (2012: 386-7n).

then be in the internal features of outcomes. I shall suggest that it is found in an *imprecision* of the relevant internal features that renders them inherently unquantifiable. This implies that the Internal View cannot conform to the numerical model Temkin assumes.

I don't believe, however, that the Internal View gives a complete account of the value of outcomes. This view captures only the *intrinsic* value of outcomes. As some of Temkin's examples show—in my opinion, Progressive Disease-Third Version (2012: 441-5) is especially persuasive—to decide whether one outcome is better all things considered than another outcome, sometimes it is not enough to consider the intrinsic value of these outcomes. It is also necessary to take into account various *relations* between these outcomes, such as relations of identity between people in these outcomes. (As we have seen, the fact that the people in A+ are identical to the people in B, but not to the people in A may be thought to make the ideal of equality relevant only to a comparison of the first two outcomes.) However, in many cases it *is* enough to consider the intrinsic value of outcomes in order to establish which one is better all things considered. My claim is that even in such cases the relation of being better than all things considered isn't transitive. If so, the reasons for the nontransitivity can't lie solely in variations of the evaluative factors that Temkin regards as distinctive of the Comparative View.

To my mind, a case in point is the series of outcomes leading to the Repugnant Conclusion. This conclusion can be reached via the sort of Mere Addition case considered above, but it can arise simply from a spectrum of cases in which the level of well-being, the same for everybody, in each outcome is slightly lower than it is in the preceding outcome, but the number of individuals is twice (or several times) as high. Here it is reasonable to hold that the second outcome, B, is better all things considered than the first, A, because the increase in quantity (i.e. the number of subjects receiving well-being), outweighs the loss of quality, the lowering of their level of well-being. Thus, when we think about such outcomes, we are inclined to adopt what Temkin calls an *additive-aggregationist* position.

To continue the descent towards the Repugnant Conclusion: a third outcome, C, in which the level of well-being is slightly lower than in the second outcome, but in which there are twice as many people as in B is similarly better than B, and so on. However, at least when the level of well-being becomes so low that it is barely above the neutral level, we are inclined to think that *however* many the individuals are in this outcome, Z, (as long as it's a finite number), it's worse all things considered than an outcome in which many individuals exist at a very high level, e.g. A. Thus, we adopt an *anti-additive aggregationist* stance when the differences in quality or level of well-being become sufficiently extensive. But this is inconsistent with the belief that the

relation of being better than all things considered is transitive: if B is better all things considered than A, C is better all things considered than B ... and Z is better all things considered than Y, then, if this relation is transitive, it follows that Z is better all things considered than A.

Like Temkin, I am inclined to block the Repugnant Conclusion by rejecting transitivity of “all things considered better than”.² But I find it unsatisfactory to hold, as does the Comparative View that Temkin advocates, that this nontransitivity is based on a variation of the relevance or significance of evaluative factors in the process of comparing different outcomes of the spectrum, i.e. that the additive aggregationist and anti-additive aggregationist positions involve a variation of factors. Let’s make the matter as simple as possible by assuming that the differences in respect of level of well-being in the Repugnant Conclusion spectrum derives from differences in respect of the intensity of some sort of physical pleasure, for instance, the sensations we have when we eat the tastiest food, the second tastiest food, and so on. We’re inclined to think that when the intensity of a pleasure becomes high enough, this pleasure, if of adequate duration, is better than a pleasure whose intensity is sufficiently lower, *whatever* the duration of the latter may be. However, it isn’t easy to see how this is possible, since this difference in intensity and duration *consists* in a number of smaller, intermediate differences in these respects. How can the value of this difference be anything other than a sum of the differences in which it consists? This question is especially pressing if one adopts a numerical model of these internal factors, as does Temkin.

He claims that in such a spectrum “the relevant factors, or the significance of those factors, for comparing ‘distant’ alternatives A and Y, may differ from the relevant factors, or significance of those factors, for comparing intervening ‘adjacent’ alternatives” (2012: 224). Therefore, he claims that the value of Y might be n when Y is compared to X, but different, o , when Y is compared to A (see 2012: 229). This fits the value of $A+$ when a Comparative View is taken of the ideal of equality; then the value of $A+$ will be higher when it is compared to A than when it is compared to B, since in the latter case the badness of inequality detracts from its value. In other words, one “factor”, the factor of equality, which is relevant for the comparison $A+$ and B, is irrelevant for the comparison of $A+$ and A.

By contrast, in the Repugnant Conclusion series, irrespective of whether we compare the value of adjacent outcomes, like Y and X, or distant outcomes like Y and A, the factors seem to be *the same*, namely the intensity and

2. He also considers the possibility of avoiding the Repugnant Conclusion by a Capped Model of Utility (2012: 328ff). But this model rejects the idea that “utility is intrinsically valuable” (2012: 343) which I regard as strongly counter-intuitive.

duration of certain sensations (on my assumption above). These are internal features of these outcomes. Moreover, the difference in intensity and duration between Y and the distant A *consists* in a number of adjacent differences in these respects; so how can the value difference between Y and A be anything but *the sum* of the adjacent value differences, i.e. how is anti-additive aggregationism possible? It seems to me that, having found that variations in respect of evaluative factors explain nontransitivity in other cases, like Mere Addition, Temkin simply *postulates* that such variations are also present in the Repugnant Conclusion series because there is apparently nontransitivity. But if this is his *only* reason for postulating such variations, it isn't convincing to claim that they make up the explanation of nontransitivity, unless all alternative explanations have been excluded. Also, one would like a more detailed account—of the sort Temkin supplies as regards Mere Addition—of how evaluative variations generate nontransitivity in the series. Otherwise, the reasonable conclusion might be that the apparent nontransitivity is just that—apparent.

I shall suggest an alternative, more detailed explanation of how anti-additive aggregationism with respect to value is possible, of why bigger value differences aren't necessarily identical to sums of intermediate value differences in the spectrum, but a big enough value difference such as between A and Y can be *greater* than the sum of the intermediate value differences A-B, B-C, ... X-Y. It's made possible by an *imprecision* as regards what has intrinsic value, e.g. pleasure and pain, which is incompatible with an assignment of numerical values to their goodness or badness. In other words, we must reject numerical models of the value of pleasure which assigns a number to the value of a pleasant experience which is based on its duration and intensity. If we reject such numerical or linear models of value, the source of the nontransitivity of "all-things- considered better than" can lie in the internal features of outcomes that the Internal View recognizes.

2. AN ARGUMENT AGAINST THE TRANSITIVITY OF IDENTITY OF SUPERVENIENT PROPERTIES

It's relatively uncontroversial that the value of a thing *supervenes* on other properties of it. I shall now present an argument to the effect that the relation of *sameness* or *perfect similarity* with respect to properties that are supervenient can't be transitive. As we shall see, this argument implies that the relation of being better than all things considered isn't transitive.

According to an informal understanding of the notion of supervenience, to say that S is a supervenient property of X means that there have to be other properties of X, basic properties, B, *in virtue* or *because of* which X has S,

properties that *determine* or *explain* X's having S. I shall say no more about this notion than that I take it to imply that supervenient properties aren't logically entailed by or identical to their basic properties.³ For instance, X's having some less specific or more determinable property such as being yellow or green, or being coloured, isn't supervenient on X's being yellow because the former properties are obviously entailed by the latter. They don't seem to be genuinely supervenient on X's being yellow, since X's having them isn't ontologically anything *additional* to X's being yellow. Similarly, the weight and spatial properties of a thing don't supervene on the weight and spatial properties of its proper parts, since if one knows the weight and spatial properties of all parts, one can deduce or calculate the weight and spatial properties of the thing they constitute. Thus, the property of the whole isn't anything over and above the sum of the properties of the parts. Nor is a dispositional property like X's being brittle supervenient on X's having a certain molecular structure if being brittle *is* having this structure (identified in terms of how it responds to certain causes).

As already implied, the properties of being of positive or negative (not neutral) value, or being good or bad, are usually regarded as paradigm examples of supervenient properties. Another set of often cited examples are so-called secondary qualities, such as having a particular colour or taste. However, as has been pointed out, e.g. by Simon Blackburn (1988: 66 ff), it appears to be a matter of *linguistic competence* to know that value properties are dependent on other properties. You show that you are not in command of the terms "good" and "bad" if you think that something can be good or bad without being so in virtue of some other properties that it has. In contrast, knowing how to apply colour terms apparently doesn't entail knowledge that if something has some colour, it has it in virtue of some other properties.

This difference with respect to linguistic competence between values and secondary qualities may be the reason why the notion of supervenience was first introduced in value theory by, G. E. Moore (1922: 261).⁴ But this difference doesn't make it uncontroversial that values are supervenient in the sense that I've adopted which implies that they're ontologically distinct from subvenient properties. This claim needs to be defended both in the case of values and secondary qualities, though I shan't do so here.⁵

3. In terms of Jaegwon Kim's distinction between strong and weak supervenience (see, e.g., "Concepts of supervenience" reprinted in Kim, 1993), the essential point is that the dependence mustn't be so strong that it jeopardizes the distinctness or irreducibility of the supervenient properties.

4. The *term* was introduced by Hare (1952:145).

5. For instance, value and secondary qualities will be supervenient in the requisite sense if they're subjective in the following fashion: value is definable in terms of desire fulfilment

It follows from the informal notion of the supervenience of a property that if there's something, Y, that's perfectly similar to X in respect of the basic properties B, then X and Y are also perfectly similar in respect of the supervenient S. In other words, if there's a difference between X and Y with respect to S—if X has S but Y lacks it, or if X has S to a greater or smaller degree than Y, etc.—there must be a difference in respect of B between X and Y. Otherwise the difference in respect of S can't be explained in terms of B.

In contrast, supervenience doesn't imply that if there's no difference or perfect similarity between X and Y with respect to S, then there's no difference between them with respect to B. Differences in respect of B which are insufficient to generate differences in respect of S may well exist.

For the purposes of my argument, I needn't plunge any deeper than this into the notion of supervenience, since I've already secured the simple implication of supervenience which forms the first premise of my argument:

Simp: If S is a property of objects that supervenes on their having B, then, for all objects X, Y and Z, even if both X and Y, and Y and Z, are perfectly similar or the same with respect to S, it's logically possible that there are differences with respect to B between both X and Y, and Y and Z.

To illustrate: even if the physical stimulations X and Y are felt to be equally painful and bad to humans, and the same is true of Y and Z, it may be that there are differences between X and Y and Y and Z that are too small to be registered and transmitted by the human nervous system to the brain.⁶ This is an example of the very simplest kind of value judgement. As will emerge, there's reason to focus on such simple examples. But although this example is as simple as they come, it's controversial how it should be properly understood. There are three things whose precise relations to each other are debatable: the painfulness of a sensation (for the subject), the (intrinsic) dislike of or aversion to it (which the subject has), and the (intrinsic) badness of it (for the subject). I shall assume that a pain is bad (for the subject) because it's disliked (by the subject), and disliked because it's painful. The precise force of these "because", whether they're conceptual or contingent, doesn't matter for present purposes.

The next step in the argument is a claim about the possible differences as regards B between X and Y, and Y and Z, of which *Simp* speaks:

rather than in terms of the properties that make things valuable, and secondary qualities are properties whose nature is revealed in our experiences rather than any properties which cause these experiences.

6. I take a "stimulation" to have both a physical aspect (a cut, burn, etc.) and a psychological aspect (a sensation). Both can be said to be painful.

Add: Even if there are differences in respect of B between X and Y, and between Y and Z, neither of which are sufficient for differences in respect of S between X and Y, or between Y and Z, but X is perfectly similar to Y, and Y to Z, with respect to S, it's *logically possible* that there are differences in respect of B between X and Z that are sufficient for a difference with respect to S between X and Z.

The fact that the differences with respect to B between neither X and Y nor Y and Z are sufficient for there to be a difference with respect to S between X and Y, or Y and Z, is surely compatible with there being *another* difference in respect of B between X and Z which *is* sufficient to manifest itself in a difference with respect to S between X and Z. Since B and S are distinct properties, the sufficiency in question is contingent, e.g. causal. But, evidently, the fact that neither the difference in respect B between X and Y nor between Y and Z is contingently sufficient for a difference in respect of S between them can't logically entail that the difference in respect of B between X and Z—which may be twice as big as either of the two other differences—isn't contingently sufficient for a difference as regards S between X and Z. For instance, the following is clearly logically possible: the difference between X and Y and between Y and Z with respect to B (e.g. the pain-producing properties) is each one unit, but the difference in this respect between X and Z is two units, and a difference of two units is minimally sufficient to give rise to a difference as regards S for the subjects in question.

This additive possibility is one reason for the name of the second step of the argument. Another reason is that it's an additional premise, supplying the link between supervenience and transitivity. This link comes out in the third step:

Trans: If *Add* is true, it must be false that the relation of perfect similarity or sameness with respect to S is transitive, i.e. it must be false that it's a *logically necessary* truth that if X and Y, and Y and Z, are perfectly similar or the same with respect to S, then X and Z are perfectly similar or the same with respect to S.

If *Add* is true, it must be logically possible that there be a difference with respect to S between X and Z, though there's no such a difference between X and Y, or between Y and Z, for, as we have seen, the latter similarities are compatible with there being a difference with respect to B between X and Z which is *sufficient* to manifest itself in a difference with respect to S. If it's logically possible that there's a difference which is sufficient for the manifestation of another difference, it must be logically possible that the second difference obtains.

Now, from *Add* and *Trans* we may validly infer by means of *modus ponens*:

Conclusion: The relation of perfect similarity or sameness with respect to a supervenient property S isn't transitive.

According to *Simp*, this is true of S because of something that follows from the fact that S is supervenient, namely that there may be differences in the subvenient properties, though there's no difference in the supervenient ones. So, the nontransitivity of perfect similarity or sameness as regards these properties follows from their supervenience.

It may *in fact* be true of some, or even all, objects X, Y and Z, that if X and Y, and Y and Z, are perfectly similar with respect to S, X and Z will also be perfectly similar in this respect. This may be because it's in fact not only the case, as the notion of supervenience implies, that if two things are perfectly similar in respect of B, they are also perfectly similar as in respect of S but, conversely, that if they are perfectly similar in respect of S, they are perfectly similar in respect of B. This possibility refutes the (implausible) claim that the relation of perfect similarity as regards supervenient properties is *intransitive*, not my claim that it is *not* transitive, or nontransitive. The fact that in some cases it's true that, if X and Y, and Y and Z, are perfectly similar as regards S, then X and Z are perfectly similar as regards S, doesn't establish that this is so as a matter of *logical necessity*, which is what the denial of the transitivity of the relation of perfect similarity denies.

It's then logically possible that, even though the difference with respect to B between X and Y, B_{xy} , is insufficient to make a difference in respect of S between X and Y, and another difference with respect to B between Y and Z, B_{yz} , is insufficient to make a difference with respect to S between Y and Z, a third difference with respect to B between X and Z, B_{xz} , is sufficient to make a difference in respect of S between X and Z. The heart of my argument is that, since this possibility statement is incompatible with the necessity statement which expresses transitivity as regards sameness with respect to S—that it's necessary that if X and Y, and Y and Z, are the same with respect to S, so are X and Z—the latter statement isn't true when the former is.

Notice that it isn't possible to argue "top-down" that since X and Y and Y and Z are identical as regards S, and identity is a transitive relation, there can't be differences as regards subvenient properties which are sufficient for a difference between X and Z as regards S. Since it's subvenient properties which determine the supervenient properties, and not the other way around, the argument has instead to be "bottom-up": because it's possible that the difference B_{xz} is sufficient for a difference between X and Z in respect of S, transitivity has to go.

3. THREE OBJECTIONS TO THE ARGUMENT

3.1. An Appeal to Relativity

Let's now consider some objections to this argument. I've assumed that the supervenient properties that X, Y and Z have are *intrinsic* properties, roughly, properties that they have independently of their relations to things external to them.⁷ Also, I take the relation of perfect similarity or sameness to be intrinsic in the sense that it holds between two relata independently of their relations to anything external to them. One objection challenges the assumption that the supervenient properties of X, Y and Z are intrinsic and claims that they're instead relative to what the object of comparison is:

Rel: How something, e.g. Y, is with respect to S depends on whether it's compared to X or Z.

Rel offers to explain how Y can be perfectly similar as regards S to both X and Z, though the latter aren't perfectly similar to each other, by claiming that Y isn't the same with respect to S when it's compared to X as when it's compared to Z.⁸ But if how Y is as regards S depends on the object to which Y is compared, how Y is as regards S can't be an intrinsic feature of Y. It has to be an extrinsic feature, which is dependent on whether X or Z is the object of comparison. *Rel* undercuts my argument against transitivity, but it does so by implying that transitivity isn't applicable to the case at hand, since this applicability presupposes that how Y is with respect to S is the same irrespective of whether it's compared to X or Z. Otherwise, there's no basis for any inference as to how X and Z are as regards S.

In defence of *Rel* it may be said that it's impossible that one and the same thing, Y's S-ness, can be perfectly similar to both X's and Z's S-ness, when X and Z are different with respect to S. But this is in fact not impossible if S is a supervenient property and how something intrinsically is with respect to S can remain the same, though the stimulation which is its cause varies within a certain range. For then the stimulation Y may lie within the same range as the stimulation X and within the same range as stimulation Z—at a point at which these ranges overlap— though X and Z do not lie in the same range.

7. In order to cater for the possibility that an intrinsic property is subjective in the sense (alluded to in footnote 5) that X's having such a property is analyzable in terms of some subjects reacting with some psychological state to X under certain conditions, one should say, alternatively, that a property is an intrinsic property of X if it's possible to *determine* that X has it by considering only X and its parts.

8. This is a point that Robinson (1972) and Jackson & Pinkerton make (1973).

Imagine, again, that it takes a difference of two units of physical stimulation for there to be any difference in respect of S (painfulness, say) and that X consists in one unit of stimulation, Y in two units and Z in three units. X is then S-in-virtue-of-one-unit, Y is S-in-virtue-of-two-units, and Z is S-in-virtue-of-three-units. By hypothesis (since a one-unit difference in stimulation is insufficient to give rise to a difference as regards S), X and Y are qualitatively identical as regards S-ness, and so are Y and Z. Hence, being S-in-virtue-of-one-unit-or-two-units expresses one kind of S-ness, and so does being S-in-virtue-of-two-units-or-three-units. But being S-in-virtue-of-one-unit-or-two-units and being S-in-virtue-of-two-units-or-three-units do not express one kind of S-ness, since X which is S-in-virtue-of-one-unit and Z which is S-in-virtue-of-three-units differ in respect of S. Nevertheless, since Y is S-in-virtue-of-two-units, and this is a common element of the two disjunctive properties, Y is both S-in-virtue-of-one-unit-or-two-units and S-in-virtue-of-two-units-or-three-units. So, Y can after all exhibit just one kind of intrinsic S-ness and still be perfectly similar in respect of intrinsic S-ness to both X and Z, though X and Z are distinct from each other as regards intrinsic S-ness. Therefore, although I can't rule out *Rel* as a possible description of the case at hand, my interpretation, according to which it violates transitivity, is also possible.

3.2. An Attempted Parallel with Primary Properties

A second line of attack in effect draws a parallel between perfect similarity in respect of a supervenient property and perfect similarity in respect of a property whose exemplifications can't be explained in terms of the exemplifications of any other kind of property. I shall call the latter properties *primary*.

Suppose that both X and Y, and Y and Z, are perfectly similar with respect to a primary property, P. This is of course compatible with their being different with respect to other primary properties, Q. But no difference in respect of Q can make it the case that X differs from Z in respect of P since, by hypothesis, an object's having P isn't supervenient on its having some other property. Hence, if there's a difference between X and Z with respect to P, this must be explained by or grounded in there being a difference in respect of *this* property, P, between X and Y or Y and Z (or both). So, if no such differences between X and Y, or Y and Z, are observed, unobservable differences as regards P must be postulated.

Imagine that one observes a difference in length (weight, etc.) between X and Z, though according to our most accurate measurements of X and Y, and of Y and Z, the members of these pairs are equally long. Then, provided

that the comparison of X and Z is correct, we're forced to conclude that there *must* be unobserved or unmeasurable differences *in length* between X and Y or Y and Z (or the particles constituting them and the spaces in-between these particles).

The present objection makes a parallel claim as regards supervenient properties:

Par: As in the case of primary properties, there must be unobservable differences in respect of S between X and Y and/or Y and Z, when there's an observable difference in respect of S between X and Z.

Derek Parfit seems to assume the existence of such unnoticeable differences as regards pain when he writes:

"I believe that someone's pain can become less painful, or less bad, by an amount too small to be noticed. Someone's pain is worse, in the sense that has moral relevance, if this person minds the pain more, or has a stronger desire that the pain cease. I believe that someone can mind his pain slightly less, or have a slightly weaker desire that his pain cease, even though he cannot notice any difference." (1984: 79).

We have assumed, with Parfit, that a sensation of pain becomes less painful "in the sense that has moral relevance", i.e. becomes less bad, when one has a weaker desire that it cease. To fail to notice that one's desire that a pain cease has become weaker then amounts to failing to notice that the pain has become less bad. Certainly, people can be suspected of sometimes making mistakes when they introspectively investigate their desires. For instance, somebody who sincerely reports not minding people of other races might be suspected of being mistaken if he's observed to avoid the company of such people. This is because the behavioural evidence contradicts his introspection.

But imagine instead that the behavioural evidence supports the introspective finding that there's no difference in respect of degree of being disliked between X and Y, and between Y and Z: the subject doesn't show any sign of choosing one member of these pairs in preference to the other. Imagine, for instance, that he's simultaneously pricked by pins in his left and right hands, and both his introspective scrutiny and his behaviour support the view that there's no difference in his aversion to the two pains. Then it seems that there's no reason to hypothesize that the subject's aversion to one pain is stronger than his aversion to the other, and, hence, that one pain is for him worse than the other than that this case is parallel to a case in which the comparison concerns some primary property (and there is transitivity).

Par's claim that there *must* be such unfelt or unnoticeable differences as regards S between X and Y, or Y and Z, and, hence, that the difference in respect of basic properties between X and Y, B_{xy} , or the difference in respect of such properties between Y and Z, B_{yz} , *must* be sufficient for a difference in respect of S because there is a difference between X and Z in respect of S, betrays a misunderstanding of the nature of supervenient properties, of their dependence on subvenient properties. For the difference in respect of S between X and Z can be explained in terms of the difference B_{xz} , of this difference being sufficient to make up a difference in respect of S between X and Z, even though each of the differences B_{xy} and B_{yz} isn't sufficient for a difference in respect of S. Therefore, a postulation of unobservable differences *in respect of S* between X and Y, and between Y and Z, isn't *necessary* to make the difference between X and Z intelligible, as it is in the case of primary properties.

Now, if it isn't necessary that there be a difference in respect of S between either X and Y, or Y and Z, when there's such a difference between X and Z, there's no transitivity with respect to S. Transitivity implies that it would be incoherent to assume that there's a difference in respect of S between X and Z, though there's no such difference between X and Y or Y and Z. But this isn't incoherent, since the difference between X and Z can be explained in terms of differences as regards the subvenient properties that, as a matter of definition, must underlie supervenient properties. The supervenience of S guarantees that there's a difference like B_{xz} to which an explanation of the difference in respect of S between X and Z can refer.

Furthermore, the postulation of such unfelt and unnoticeable differences in respect of S isn't only redundant, but impossible, if S is a subjective property, as painfulness presumably is. If the *esse* of pain is *percipi*, it *can't* have any unfelt aspects. This rules out that my considered judgement that X and Y are feeling equally painful to me can be rejected because of differences in painfulness that aren't felt by me. Analogously, it's difficult to see what sense could be given to the concept of dislike or aversion if it isn't construed as something that entails either behavioural tendencies of withdrawal or differences of feeling. But in the cases we have examined, there aren't, by hypothesis, differences of either kind.

It should be re-emphasized that I've assumed that supervenient properties are distinct from their bases. We shouldn't expect perfect similarity to be nontransitive as regards properties which are disqualified from being supervenient by this distinctness requirement, but in some instances it is. For instance, it's nontransitive as regards the properties of being yellow or green and having some colour. Imagine that X is a two-coloured object, both yellow and green. So are Y and Z. Imagine further that Y has the same yellow shade as X, but a different green shade, whereas Z has the same green shade

as Y, but a different yellow shade. Then X and Y are perfectly similar as regards both the disjunctive property of yellowness or greenness and some of their colours, and the same is true of Y and Z. But X and Z are not perfectly similar as regards either of these properties.

Some identify secondary qualities with primary qualities, e.g. having some colour with reflecting light of a certain wavelength. Likewise, according to some metaethical theories, the relation between value properties and the properties they're supposed to supervene on is entailment as in the case of the determinable properties under discussion. If this is right, secondary qualities and values won't be supervenient in my sense, and my argument doesn't apply to them. It's only if value properties and secondary properties are taken as supervenient in a sense which, like mine, implies that they're entirely distinct from their bases that my argument rules out transitive sameness with respect to them.

3.3. An Appeal to Roughness of Comparison

Comparisons concerning supervenient properties may be rough or imprecise. A comparison of value will be rough, e.g. when you compare complex and qualitatively different things which must be evaluated along several dimensions that have to be weighed against each other, especially if evaluation along some of these dimensions is to some extent subjective. Suppose, for instance, that you were to evaluate the greatness of novelists that are very different as regards style, content etc., like Joyce, Kafka and Proust. The judgement that these writers are equally great novelists is obviously rough because no precise comparison between them is possible. If you compare a long string of novelists who are only roughly comparable, you may find it impossible to distinguish between the greatness of one writer and the next and still end up with writers who are distinguishable in respect of greatness because smaller differences eventually mount up. So, it isn't to be expected that judgements of sameness under conditions of rough or imprecise comparability ensure transitivity. A final objection makes an appeal to such roughness:

Rough: Similarity in respect of S between X and Y, and Y and Z, can only be rough, or imperfect, when X and Z aren't the same with respect to S, though X and Y and Y and Z are.

However, a comparison between two simultaneous stimuli in respect of painfulness and badness can clearly be more precise than such rough comparisons as between the aforementioned novelists because it's a comparison of much simpler entities. It seems to me that it can be as precise as any

comparison of the (intrinsic) value of two things can possibly be. The upshot of such a comparison can therefore be that one stimulus is precisely and perfectly as painful and bad as the other in the sense that there's *no difference whatsoever* between them *in respect of painfulness and badness*. So, this comparison isn't rough or imprecise in the sense that there are differences in these respects that it fails to register, as a comparison of the weight of two things based on lifting them is.

It may be objected that someone with more acute pain receptors than mine could feel a difference between the stimulations X and Y (and Y and Z), though I don't. True, but this doesn't show that there's any difference between the pains I am feeling when I'm stimulated by X and Y, since this subject is feeling two kinds of pain where I'm feeling only one kind. Consider a creature whose pain receptors are as acute as you like. Realistically, there will still be minute differences in physical stimulation which aren't distinguishable by this creature, but suppose there isn't. Then this creature would never feel any difference in respect of painfulness between X and Z when it feels no such difference between X and Y, and between Y and Z. But, as long as the supervenient and subvenient qualities are distinct, this would still not be so as a matter of necessity.

For the same reason, there being unfelt differences with respect to primary properties doesn't imply that there are after all unfelt differences *as regards the supervenient properties*, or that the comparison with respect to the latter properties isn't as precise as it could be. This is something that distinguishes the argument that I've been running from Sorites arguments in which unnoticeable or negligible differences in respect of a feature add up to a difference that is noticed or significant. If you remove one grain from a heap, you will still have a heap, but —whether or not you notice it— you won't have something which is *the same* in respect of "heapiness", as *much* of a heap. You will have something which is *less* of a heap, and if you go on removing grains, you will eventually have something which isn't a heap at all. This is because something's being a heap is *constituted* by a number of grains standing in certain spatial relations to each other. The former isn't a distinct property which supervenes on the latter in my sense of the term, as the painfulness supervenes on certain physical stimulation. So, it can't remain the same when the constitutive elements undergo changes.

I've considered the strategy of supporting the roughness of which *Rough* speaks by (a) taking supervenient properties to be analogous to primary properties. But it could also be supported by (b) claiming that the relation of perfect similarity or sameness is, *by necessity or definition*, transitive. It could be claimed that the comparison between, e.g. the stimuli Y and Z in respect of painfulness isn't as precise as possible if only Y and Z themselves are considered, that this requires comparing them to other things, such as

X. So, Y and Z would be perfectly similar in respect of S only if, for X and all other things, Y and Z are perfectly similar in respect of S to X and other things (see Goodman 1951: 221). Failure of perfect similarity to other things is now no longer seen merely as *evidence* of there being undetected differences and, thus, no perfect similarity between Y and Z, but as something that *entails* that there's no perfect similarity between Y and Z.

But, apart from the fact that this way of defending *Rough* is blatantly question-begging in the present context, it's at odds with the intuition that perfect similarity in intrinsic respects between two entities, like Y and Z, is an *intrinsic* relation between them such that it's possible to determine whether it holds between them by examining only how Y and Z are as regards the relevant respects, and nothing external to them like X. This is a way to understand the relation in the case of intrinsic primary properties without getting nontransitivity as an inescapable result.

Turning to another way of buttressing (b), it may be true as a matter of definition that if X and Y are perfectly similar in respect of a feature, such as S, there's a feature S such that X has S if and only if Y has S. But it can be seen from my reply to *Rel* that it doesn't follow from this claim that the notion of perfect similarity with respect to a feature is by definition transitive. I imagined the S-ness that both X and Y have to be S-ness-in-virtue-of-one-or-two-units. But Y, which was imagined to be S-in-virtue-of-two-units, also has the property of S-ness-in-virtue-of-two-or-three-units, and we saw that this could be a distinct kind of S-ness from S-ness-in-virtue-of-one-or-two-units. Y and Z could be perfectly similar with respect to this S-ness in the sense that Y has this S-ness if and only Z has it. This obviously does not entail that there is any kind of S-ness such that X has it if and only if Z has it. Consequently, we should reject the claim that the proposed definition *entails* that the relation of perfect similarity or sameness as regards a property is transitive. This is rather an implication that the definition has when applied to properties that I've called primary, since the phenomenon that I've just described is impossible in the case of primary properties because there are no other properties underlying them.

However, suppose that you're feeling the sensations X, Y, and Z simultaneously. Surely, you can't simultaneously compare X and Y, Y and Z, and find them the same, and X and Z and find them different. I agree. The most accurate comparison between two sensations requires undivided attention to them, but when you try to execute the three comparisons at the same time, none of them gets undivided attention. This makes it likely that a minute difference between X and Z won't be noticed. So, all three sensations would come out as identical. If the series of sensations had consisted not in three, but in ten sensations, the difference in intensity between the end-points could be greater, but then the distraction would also be greater, so bigger dif-

ferences may escape unnoticed. I don't think, then, that my account carries any implausible phenomenological implications. It doesn't imply that the relations it describes are noticeable in *all* circumstances. Indeed, it's possible that they aren't noticeable in *any* circumstances – this might be true if the pains are so intense that we can't ever concentrate on the task of making accurate comparisons.

3.4. The Nontransitivity of Better/Worse Than

I see no other way to resist my argument to the effect that the relation of identity or perfect similarity as regards properties that are supervenient isn't transitive. I've assumed, rather than argued, that value properties and secondary qualities are supervenient in the sense explained at the outset. If no properties should turn out to be supervenient in this sense because all candidates are entailed by or identical to their bases, my argument will be less interesting, but I'm inclined to think that this isn't anything I need to fear.

Now if values are supervenient, it's easy to see how my argument could be extended into an argument against the transitivity of the relations of being better/worse than all things considered.⁹ Imagine, for instance, that the painful stimulation Y is slightly shorter than X, though there's no difference in the felt intensity of the pain. Then Y is better than X all things considered (assuming that intensity and duration are the only relevant factors). If the same is true of Y and Z, Z will be better than Y all things considered. Still, it might be that Z isn't better than X all things considered because Z is felt to be more intense than X, and this difference is judged to outweigh the longer duration of X.

To take an example more similar to Temkin's spectrum cases, imagine that X is minimally more painful than Y, but that Y is markedly longer than X. The same goes for Y and Z. Then Y may be worse than X all told, and Z may be worse than Y all told because the greater difference in duration outweighs the smaller difference in intensity. Yet, Z may be better than X all told because, due to unmanifested differences in the subvenient properties, the dif-

9. John Broome claims that it's an analytic or conceptual truth that comparative relations are transitive (2004: 50-1). But, to return to the two-coloured objects considered earlier, suppose that the green of Y is greener than the green of X, but its yellow is paler. Likewise, the yellow of Z is yellower than the yellow of Y; it's in fact as yellow as the yellow of X, just as its green is as green as the green of X. Then Y is yellower or greener (or more saturated in respect of some colour) than X, and Z is yellower or greener than Y, but Z isn't yellower or greener than X. So, the comparative relation of being yellower or greener (or of being more saturated in respect of some colour) isn't transitive. Hence, Broome's claim is false. Temkin gives the similar counterexample of "larger than" defined as "heavier or taller than" (2012: 164).

ference in intensity between X and Z may exceed the sum of the differences in this respect between X and Y, and Y and Z. Therefore, it is possible that the difference in intensity between X and Z in Z's favour overpowers the difference in duration in X's favour. In the case of both examples, nontransitivity may come out as more plausible if the series are made longer.¹⁰

This nontransitivity isn't due to any variation in respect of what factors are relevant; they are throughout the intensity and duration of sensations. It has to do with a variation of their "significance" if this refers to the shift from an additive aggregationist approach to an anti-additive aggregationist approach, but the explanation of this shift and nontransitivity lies in the imprecision of the intensity of pain relative to the underlying physical stimulation. The intensity and duration of sensations are internal features of pains; so, a value that depends on them is captured by the Internal Aspects View.

Rethinking the Good is a first-rate intellectual achievement. However, my conclusion is that, despite all his ingenuity and thoroughness, Temkin has missed that there's an Internal Aspects View of the intrinsic value of outcomes – a value which in many instances determines their value all things considered – that implies that values aren't quantifiable in a way that guarantees transitivity. Obviously, we can't assign any number to the value of sensations which is based on their intensity and duration, since the number of this value of Y would have to be the same as the numbers assigned to X and Z, though these would have to be different from each other. This Internal View seems to me to provide a more informative explanation of some simple spectrum cases than does Temkin's Comparative View.

But since I've agreed that the Comparative View can provide a rationale for nontransitivity in some cases, it may be, for all I've said, that some other spectrum cases involving more complex value judgements are among these cases. However, although I've indicated that Temkin's Comparative account of why we adopt an anti-additive stance when we compare distant outcomes in a spectrum needs to be filled out, I don't think it can be filled out in the terms I've supplied without impugning the distinction between Comparative and Internal Views. Therefore, my conclusion is that Temkin has to concede that the Internal View provides the basis for the nontransitivity of better than all things considered in some cases and, thus, is capable of providing such a basis.¹¹

10. Stuart Rachels has also presented such arguments (1998) and (2001).

11. Sections II and III of this paper have been presented at Pompeu Fabra University, Rutgers University, University of Massachusetts at Amherst, University of Gothenburg, University of Copenhagen, The Moral Philosophy Seminar, University of Oxford, and University of York. Many thanks to these audiences for valuable comments. But a special thanks to Larry Temkin for many stimulating discussions of this topic through many years.

BIBLIOGRAPHY

- Blackburn, S., 1988: "Supervenience Revisted", in G. Sayre-McCord (ed.): *Essays on Moral Realism*, Ithaca, CT: Cornell U. P.
- Broome, J., 2004: *Weighing Lives*, Oxford: Clarendon Press of Oxford University Press.
- Goodman, N., 1951: *The Structure of Appearance*, Cambridge, MA: Harvard University Press.
- Hare, R. M., 1952: *The Language of Morals*, Oxford: Oxford University Press.
- Jackson, F., and Pinkerton, R. J., 1973: "On An Argument Against Sensory Items", *Mind* 92: 269-72.
- Kim, J., 1993: *Supervenience and Mind*, Cambridge: Cambridge U. P.
- Moore, G. E., 1922: *Philosophical Studies*, London: Routledge & Kegan Paul.
- Parfit, D., 1984: *Reasons and Persons*, Oxford: Clarendon Press.
- Rachels, S., 1998: "Counterexamples to the Transitivity of *Better Than*", *Australasian Journal of Philosophy* 76: 71-83.
- 2001: "A Set of Solutions to Parfit's Problems", *Noûs* 35: 214-35.
- Robinson, H., 1972: "Professor Armstrong on 'Non-Physical Sensory Items'", *Mind* 91: 84-7.
- Temkin, L. S., 2012: *Rethinking the Good*, New York: Oxford University Press.

Intransitivity and the Internal Aspects View

LARRY TEMKIN

Rutgers University

Abstract

This article responds to Ingmar Persson's article "Internal or External Grounds for the Nontransitivity of 'Better/Worse than'". In his article, Persson argues in favor of an account of supervenience that would be compatible with both an *Internal Aspects View*, and the *nontransitivity* of the "better or worse than" relations. This article points out that the Internal Aspects View that Persson favors would fail to capture many features of practical reasoning that most advocates of an Internal Aspects View favor, and that the version of the Internal Aspects View that I discuss in *Rethinking the Good* does capture. I note, however, that Persson's view would not only be compatible with my book's main claims and arguments, it would substantially buttress my results. Accordingly, I would welcome it if Persson could successfully develop and defend his view. Unfortunately, however, my article raises a number of worries about Persson's view. I consider various different ways of understanding Persson's position, and argue that none of them ultimately succeed in establishing a plausible version of a genuinely Internal Aspects View that would be compatible with the nontransitivity of the "better or worse than" relations. I acknowledge, that if Persson can ultimately make good on his claims, he will have made a substantial contribution to our understanding of the good and the nature of ideals. However, as matters now stand, I am not moved by his arguments to revise the claims I made in *Rethinking the Good*, correlating the nontransitivity of the "better or worse than" relations with the Essentially Comparative View, rather than the Internal Aspects View.

Keywords: Transitivity, Nontransitivity, Internal Aspects View, Essentially Comparative View, Practical Reasoning, Better than, Supervenience.

I'd like to thank Ingmar Persson for his response to *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning* (Temkin, 2012). I have long admired Persson, and I have learned much from him over the years. Persson's work typically displays a rare combination of insight, good sense,

and importance. Not only do I usually find his claims interesting and plausible, I usually find myself in agreement with them. I confess, however, that while I find his central claims in “Internal or External Grounds for the Non-transitivity of ‘Better/Worse than’” (Persson, 2014) interesting and important, I don’t find them plausible. Indeed, I’m not really sure how to make sense of them. In this article, I’ll mainly try to show why I find his claims puzzling and unconvincing.

Perhaps Persson will be able to adequately answer my concerns in a way that will give his claims the clarity, plausibility, and defensibility typical of his work. If so, I believe he will have made a significant contribution to our understanding of the good and the nature of ideals. But as matters now stand, I am not moved by his remarks to revise the claims I made in *Rethinking the Good*.

1.

Before presenting Persson’s central claims, and my response to them, it will be useful to start with a brief recapitulation of some of *Rethinking the Good*’s key claims. I do this both as useful background for Persson’s position and, as importantly, to illustrate that I could, in principle, accept everything Persson contends without significant revision of my own views. Indeed, if correct, Persson will have provided a new and important argument which not only fits comfortably with my larger views, but which, in fact, provides independent support and further vindication of those views.

In my book, I noted that many people make certain standard assumptions about practical reasoning. For example, most people assume various *Axioms of Transitivity*, believing, for instance, that the “all-things-considered better than”, the “all-things-considered equally as good as”, and the “all-things-considered at least as good as” relations are all transitive. Most people also assume an *Independence of Irrelevant Alternatives Principle*, believing that if one wants to know how any two outcomes O_1 and O_2 compare, it is sufficient to compare them *directly*. On such a view, how O_1 and O_2 compare to *each other* all things considered, won’t depend in any way on how either or both compare to some third outcome O_3 , or some alternative set of outcomes, O_k to O_n . Similarly, most people assume a *Principle of Like Comparability for Equivalents*, believing that if two outcomes, O_1 and O_2 , are equally good, then however O_1 compares to any third outcome O_3 , that is precisely how O_2 will compare to O_3 .

I showed that there is *one* way of thinking about ideals, which I called the *Internal Aspects View*, which had great intuitive plausibility and which, if true, would explain why each of the above principles held. More specifi-

cally, I gave a particular characterization of the Internal Aspects View which would account for such principles, one according to which the goodness of each outcome would depend solely on in the internal features of that outcome, and where an outcome's goodness could be accurately represented by a number or range of numbers on the real number line (Temkin, 2012: sec. 11.3).

However, in my book I also pointed out that there is an alternative way of thinking about ideals, which I called the *Essentially Comparative View*, that also has great intuitive plausibility. On this view, the factors that are relevant and significant for assessing an outcome's goodness may vary depending on the alternative outcome(s) with which it is compared. I argued that many of the ideals people care most about, including a *Narrow Person-Affecting View*, the *Pareto Principle*, and particularly plausible versions of *Maximin* and *Utility*, are best captured by an Essentially Comparative View of ideals rather than an Internal Aspects View (Temkin, 2012: ch. 12). I then argued that given the nature and structure of Essentially Comparative ideals, many of the common assumptions about the nature of practical reasoning may fail to hold or apply across different sets of outcomes, including the various Axioms of Transitivity, the Independence of Irrelevant Alternatives Principle, and the Principle of Like Comparability for Equivalentents.

In my book, I *also* pointed out that many important principles are limited in scope, in the sense that they are thought to be relevant and significant for comparing certain outcomes but not others. I noted that this is true of the Pareto Principle as it is commonly interpreted (Temkin, 2012: sec. 12.5), it is true of John Broome's *Principle of Personal Good*, as he presents it in *Weighing Goods* (Broome, 1991: sec. 8.1), and it is true of John Rawls's two principles of justice as he presents them in *A Theory of Justice* (Rawls, 1971: 63). I then showed that whenever it is true for such a principle that there can be three outcomes, O_1 , O_2 , and O_3 , such that the principle would apply when comparing O_1 and O_2 , but would not apply when comparing O_1 and O_3 , then such "limited scope" principles are Essentially Comparative in the sense I am employing that notion, and this opens up the possibility that the common assumptions about practical reasoning discussed above, including the Axioms of Transitivity, may fail to hold or apply across different sets of outcomes.

I suggested that one important set of cases where principles that were limited in scope came into play was in my Spectrum Arguments; where one such argument involved a Spectrum of outcomes where the first outcome involved a very long life with 15 mosquito bites per month and two years of excruciating pain, the second outcome involved a very long life with 15 mosquito bites per month and *four* years of pain *almost* as bad as that obtaining in the first outcome, the third outcome involved a very long life with

15 mosquito bites per month and *eight* years of pain *almost* as bad as that obtaining in the second outcome, and so on, where the last outcome of the Spectrum merely involved a very long life with 16 mosquito bites per month but no torture. I claimed that (1) in accordance with a position I called the *First Standard View*, which reflected an *additive-aggregationist* approach to comparing outcomes, most people would judge that for each adjacent pair of outcomes along the Spectrum, n and $n + 1$, the earlier outcome, n , would be *better* than the later outcome, $n + 1$, all things considered, that (2) in accordance with a position I called the *Second Standard View*, which reflected an *anti-additive-aggregationist* approach to comparing outcomes, most people would judge that the Spectrum's first member was *worse* than its last member, all things considered, and (3) that together, these plausible and widely-held judgments are incompatible with the Axiom of Transitivity for the "all-things-considered better than" relation (Temkin, 2012: ch. 5; and Temkin, 2014: sec. 1).

A key part of my analysis of what is going on in my Spectrum Arguments was to emphasize that the First and Second Standard View are both limited in scope, so that we regard the First Standard View as relevant and significant for comparing adjacent outcomes along my Spectrum, but not for comparing outcomes at opposite ends of the Spectrum, and vice versa with respect to the Second Standard View. This implies that most people are implicitly relying on an Essentially Comparative View of ideals in making the judgments they do regarding my Spectrum's outcomes, rather than an Internal Aspects View, and I claimed that this accounts for why the Axiom of Transitivity for the "betterness" relation fails, or fails to apply, across the different outcomes in my Spectrum cases.

Having distinguished between the Internal Aspects View and the Essentially Comparative View in the way that I did, I acknowledged in a note that there may be alternative ways of thinking about an Internal Aspects View that might allow for the non-transitivity of the "all-things-considered better than" relation (Temkin, 2012: ch. 11, note 32). However, I don't pursue this as, in fact, I believe that the best explanation of the various cases where the Axioms of Transitivity seem questionable lies in our implicitly accepting an Essentially Comparative View in thinking about those cases. Similarly, my motivation for focusing on the version of the Internal Aspects View that I did is that I think such a view is intuitively plausible, widely assumed in many contexts, and would account for many standard assumptions about practical reasoning, including the various Axioms of Transitivity, the Independence of Irrelevant Alternatives Principle, and the Principle of Like Comparability for Equivalents.

Now Persson accepts that certain ideals people attach great weight to in assessing outcomes are Essentially Comparative. He also accepts that many

ideals people value may be limited in scope, in a way that supports an Essentially Comparative View of ideals. So, he is prepared to grant that in a range of cases, people may be committed to a set of judgments that are incompatible with the Axioms of Transitivity because of essentially comparative considerations. Furthermore, he doesn't deny the intuitive appeal of an Internal Aspects View of the sort I characterized, nor does he deny that, if true, my version of the Internal Aspects View would support, and explain, many of the standard assumptions about practical rationality that I discuss. However, Persson is keen to defend the possibility that I broached in a note. More specifically, he believes that in the case of my Spectrum Arguments, we can explain the non-transitivity of the goodness of the Spectrum's outcomes solely on Internal Aspects grounds. That is, Persson believes the non-transitivity of most people's judgments about the goodness of the different outcomes in my Spectrum cases can be explained in a way that is wholly consistent with the view that, all-things-considered, an outcome's goodness depends *solely* its internal features. Moreover, importantly, Persson believes that unlike some of the other cases of non-transitive goodness that I discuss, the *best* explanation of the non-transitivity of outcome goodness in my Spectrum cases will rest on Internal Aspects grounds rather than Essentially Comparative grounds.

Naturally, in developing his position, Persson is committed to an alternative version of the Internal Aspects View than the one I offered in my book, since my version entails the various Axioms of Transitivity. This is partly what makes Persson's suggestion so intriguing and important. If the *best* way of understanding the Internal Aspects View differs from the one I offered, in that it *also* supports the non-transitive "all-things-considered better than" judgments that most people make about my Spectrum Arguments, this will make it even harder to deny my book's central conclusion that we need to significantly revise our understanding of the good, moral ideals, and the nature of practical reasoning.

2.

In light of the foregoing, it should be clear that I would welcome the success of Persson's project. Indeed, I would regard it as a friendly amendment that is clearly within the spirit of my own views, and one which would add significant weight to my book's main claims. But despite that, I am not persuaded that Persson has provided a better account of people's judgments in my Spectrum Arguments than my own. Indeed, to be completely honest, I don't even *understand* Persson's view, finding his claims about the Internal Aspects View and how it is supposed to support the rejection of the various Axiom of Transitivity deeply puzzling. In what follows, I will present my main worries about his position.

Persson's account of how best to understand what is going on in Spectrum Arguments rests on his views about supervenience. He assumes that any given supervenient property, V , could supervene on two distinct bases. Thus, it *could* be the case that a given base, B_1 , gave rise to a given supervenient property, V_1 , and that a slightly, or even wholly, different base, B_2 , could give rise to a *perfectly similar* supervenient property (Persson, 2014: p. 7 of his draft). I accept this view, which might be expressed in several ways. One might say that B_1 gives rise to V_1 , and that B_2 gives rise to V_2 , where V_1 and V_2 are *qualitatively* indistinguishable even if they are numerically distinct. In this case, V_1 and V_2 might be thought of as two tokens of the same type, perhaps type V . If the supervenient properties are values, which is the class Persson is concerned with, then we can express this by saying that V_1 and V_2 have *exactly the same* value, which would then mean that B_1 and B_2 were *exactly equally as good* as each other. Another way of expressing the same idea is simply to say that two partially or wholly *distinct* bases, B_1 and B_2 can give rise to the very same supervenient value, V_1 . As the latter way of putting the point is simpler for purposes of exposition, that is how I'll often put it in what follows.

If this is right, then Persson contends that we should accept his principle

Simp: If S is a property of objects that supervenes upon their having B , then, for all objects X , Y , and Z , even if both X and Y , and Y and Z , are perfectly similar or the same with respect to S , it's logically possible that there are differences with respect to B between both X and Y , and Y and Z (Persson, 2014: [p. 7 of his draft]).

As stated, there is every reason to accept principle *Simp*. But this is because if, as seems plausible, two distinct base objects could give rise to the very same supervenient property (or "perfectly similar" supervenient properties), then presumably three distinct base objects could also give rise to the very same supervenient property (or "perfectly similar" supervenient properties). So, where the supervenient property is a value, V_1 , *Simp* will be true as long as there could be three distinct base objects, B_1 , B_2 , and B_3 , each of which gave rise to V_1 , and there is good reason to accept that possibility.

However, Persson claims something much stronger, and more controversial, than what I readily grant regarding principle *Simp*. He contends that the differences between the bases of objects X and Y , and the bases of objects Y and Z , *could* be such that we should accept his principle

Add: Even if there are differences in respect of B between X and Y , and between Y and Z , neither of which are sufficient for differences in respect of S between X and Y , or between Y and Z , but X is perfectly similar to Y , and

Y to Z, with respect to S, it's *logically possible* that there are differences in respect of B between X and Z that are sufficient for a difference with respect to S between X and Z (Persson, 2014: ?? [p. 8 of his draft]).

Now I can see how principle Add could be true on an Essentially Comparative View of ideals. After all, on such a view it could be the case that the factors that are relevant and significant for comparing outcomes X and Z, might be different from the factors that are relevant and significant for comparing outcomes X and Y, or outcomes Y and Z. (Here, and in what follows, I have put Persson's views in terms of "outcomes" rather than "objects". This does not affect the substance of his views or my claims.) Hence, as principle Add contends, on the Essentially Comparative View, it could well be the case that the supervenient values of X and Y might be the *same* when *they* are compared, and the supervenient values of Y and Z might be the *same* when *they* are compared, and yet the supervenient values of X and Z might *not* be the same when *they* are compared. But I fail to see how principle Add can be made coherent on an Internal Aspects View, where the goodness of a given outcome depends *solely* in the internal features of that outcome.

Regarding the kind of situation principle Add is supposed to be addressing, Persson seems to believe that the two bases corresponding to outcomes X and Z, call them B_x and B_z , differ sufficiently that they would give rise to supervenient properties that were *not* perfectly similar or the same. Since we are interested in the case where the supervenient properties are values, let's say that the base properties of outcome X, B_x give rise to, or account for, the value of outcome X. I shall represent this as " $B_x \rightarrow V_x$ ". Further, suppose that, however the notion of value is ultimately understood, the letter K represents the value of outcome X. I will represent this as " $V_x = K$ ". We can then use the notation " $B_x \rightarrow K$ ", to represent the fact that the bases of value in outcome X that determine X's value are such, or make it the case, that outcome X's value is K . We can then similarly write that " $B_z \rightarrow V_z$ ", " $V_z = M$ " (where the letter M represents the value of outcome Z), and hence " $B_z \rightarrow M$ ". By hypothesis, in the cases covered by principle Add, $K \neq M$, since, by hypothesis, the base objects of outcomes X and Z differ sufficiently that the values for outcomes X and Z differ.

Now, as noted previously, on an Internal Aspects View, the value of an outcome depends solely on the internal features of that outcome. So, the value of outcome Y will supervene solely on the relevant internal features of Y that constitute the base, B_y , for Y's value. Assume, in accordance with principle Add, that while B_y is distinct from B_x , they both give rise to perfectly similar, or the same, values. I have already granted, in accepting principle Simp, that this might be the case. This means that $V_x = V_y$, and thus, $B_y \rightarrow K$.

But since, by hypothesis, $K \neq M$, it follows that it is *not* the case that $B_y \rightarrow M$, and so, contrary to principle Add, outcome Y's value will *not* be the same as outcome Z's value.

Alternatively, suppose that although the bases of value for outcomes Y and Z differ, they give rise to the *same* values for the two outcomes (as might be the case, in accordance with Simp). Since, by hypothesis, $B_z \rightarrow M$, it follows that it will also be the case that $B_y \rightarrow M$. But in that case it could not *also* be the case that $B_y \rightarrow K$, since, by hypothesis, $K \neq M$.

Putting the preceding together, since, on the Internal Aspects View, an outcome's goodness depends *solely* on the internal features of that outcome, it seems clear that on that view the object bases for Y's value should give rise to, or account for, *exactly the same* value for Y *whatever* alternative it is compared with. Accordingly, given that X's value, K, is *different* from Z's value, M, it seems clear that, on the Internal Aspects View, Y's value could be equal to X's, *or* it could be equal to Z's, but it could not be equal to both! More specifically, if, in fact, Y's internal features are such that $B_y \rightarrow K$, then, indeed, X and Y will have the same value, but Y and Z will not; while if, on the other hand, Y's internal features are such that $B_y \rightarrow M$, then, indeed, Y and Z will have the same value, but X and Y will not. It seems, then, that if we adopt an Internal Aspects View, we should reject Persson's principle Add.

Why does Persson think otherwise? It isn't clear. Perhaps Persson has something like the following picture in mind. The value bases for outcome X, B_x , determine X's value, K, the value bases for outcome Y, B_y , determine Y's value, L, and the value bases for outcome Z, B_z , determine Z's value, M. So, on the notation used above, $B_x \rightarrow K$, $B_y \rightarrow L$, and $B_z \rightarrow M$. Now it might be that K and L are so "close", that we can't distinguish them intuitively or phenomenologically. In that case, we might well regard them as "perfectly similar" or "the same". Likewise, it might be that L and M are so "close", that we can't distinguish them intuitively or phenomenologically. In that case, too, we might well regard them as "perfectly similar" or "the same". But it is perfectly consistent with those two facts that K and M are sufficiently far apart that we *can* distinguish them, and so rightly recognize them as different. Such a picture might account for any intuitive appeal that principle Add might have, even on an Internal Aspects View. Unfortunately, however, it would not justify or vindicate Add.

I have three related worries about the picture in question. My first worry is that such reasoning is reminiscent of familiar arguments for the intransitivity of the *indifference* relation concerning alternatives involving vagueness or imperceptibly small differences. It is well known that presented with three alternatives A, B, and C, two at a time, many people might be indifferent between A and B because the differences between them are imperceptibly small, and they might similarly be indifferent between B and C because

the differences between them are imperceptibly small, and yet they may not be indifferent between A and C. This is because *together* the imperceptibly small differences between A and B, and between B and C, might add up to a difference between A and C that is large enough to be perceptible and is one about which they would be concerned. But cases of this sort are like the notorious Sorites Paradoxes, such as those purporting to show that a heap of sand is the same as a single grain, or that hairiness is the same as baldness.

Now there is much to be said about standard Sorites Paradoxes, but here I shall simply note that in my book I argued that my Spectrum Arguments are not related to the standard Sorites Paradoxes, and Persson agrees with me about that (Temkin, 2012: sec. 9.2). My Spectrum Arguments do not rely on vagueness, nor do they trade on a series of imperceptible differences which together add up to a perceptible difference. Rather, my Spectrum Arguments rely on differences of quality and number which are *clearly* perceptible, and which seemingly combine in *one* way for making certain comparisons, but in a *different* way for making other comparisons.

Specifically, as noted previously, most people follow (something like) the additive-aggregationist approach of the First Standard View for comparing the outcomes of my Spectrum that are adjacent to each other. This generates a *clear* ranking between such outcomes where the “earlier” outcome is *better* than the “later” outcome, so that most would clearly prefer the former to the latter and would *not* be indifferent between them. Likewise, most people follow (something like) the anti-additive-aggregationist approach of the Second Standard View for comparing those outcomes that are at the opposite ends of my Spectrum. This *also* generates a *clear* ranking between such outcomes, so that most would clearly judge the last outcome as *better* than the first, and so would *not* be indifferent between them. But then, I submit that the proper explanation of what is going on with my Spectrum Arguments, and the root of their challenge to the transitivity of the “betterness” relation, has everything to do with the Essentially Comparative View of ideals, and nothing to do with the vagueness or accumulation of imperceptible differences which underlies the standard Sorites Paradoxes and which accounts for the intransitivity of the “indifference” relation in such contexts. Thus, if, as Persson claims, he is hoping to offer a better explanation of the Spectrum Arguments than the one I offered—one that is compatible with both the Internal Aspects View and the rejection of the transitivity of the “betterness” relation—it won’t do for him to rely on the sort of picture sketched above. As indicated, that picture mirrors the standard Sorities Paradoxes, but it does not mirror what is going on in the Spectrum Arguments.

Second, even if one *claims* that the values of outcomes X and Y are so close as to be indistinguishable, and likewise that the values of outcomes Y and Z are so close as to be indistinguishable, there is good reason to believe

that there are circumstances in which we'd be able to distinguish between at least one of the two sets of values, K and L, or L and M. More particularly, even if it were true that were we confronted *only* with outcomes X and Y, we might discern no difference in value between them, and were we confronted *only* with outcomes Y and Z, we might discern no difference in value between them, it seems likely that if we were confronted with *all three alternatives at once* we *would* discern a difference between at least two of the supposedly "indistinguishable" values.

Suppose, for example, that we were presented with all three outcomes, X, Y, and Z, at the same time. Suppose, as before, that on the Internal Aspect View, $B_x \rightarrow K$, and $B_z \rightarrow M$, where the difference in value between K and M was sufficiently great that we clearly recognized X's value to be different than Z's. In that context, as we were considering all three outcomes at once, how would we assess Y's value? It seems there are only three possibilities here that we need to consider. It might be that in that context, Y's internal features were such that $B_y \rightarrow K$. If that were so, then X and Y would have the same value, but Y and Z would not, and so principle Add would not apply. Alternatively, it might be that in that context, Y's internal features were such that $B_y \rightarrow M$. If that were so, then Y and Z would have the same value, but X and Y would not, and so once again principle Add would not apply. Or, it might be that in that context Y's internal features were such that $B_y \rightarrow L$, where L was an intermediate value between K and M. In that case, Y's value, L, will either be clearly distinguishable from one or both of K and M, or it will not. If it *is* clearly distinguishable from one or both of K and M, then once again principle Add would not apply. But similarly, if it is *not* clearly distinguishable from both K and M, then in that context we will have good reason to be confident that Y's value is *not perfectly* similar to, or the *same* as, X or Z's values, since, by hypothesis, those values, K and M, *are* clearly distinguishable from each other. Hence, again, principle Add won't apply.

In sum, for any three outcomes to which we *might* have thought principle Add would be applicable, if we considered those outcomes two at a time, we can see that principle Add would *not* apply to those three outcomes if we considered all three of them at once. Insofar as the latter result seems firmly grounded, and I believe it is, this suggests one of two appropriate responses to the initial judgment. First, we might decide that the initial judgment that principle Add applied to the three outcomes when they were considered two at a time was mistaken. In essence, we might conclude that the judgment in question was akin to a perceptual illusion, which is only revealed as such when we consider all three outcomes at once. This might be like the predicament of someone confronting the famous Muller-Lyer illusion, who was convinced when looking at two side-by-side line segments, alone, that the one with the "inward" pointing arrows ($\leftarrow\rightarrow$) was longer than the one with the "outward" pointing arrows ($\rightarrow\leftarrow$), until a ruler was placed between them revealing that,

in fact, the two line segments were the exact same length. Alternatively, we might retain our conviction in our initial judgment, that X and Y really do have the *same* value, K, when they are compared with each other, and that Y and Z really do have the *same* value, M, when they are compared with each other, but that X and Z have *different* values, K and M, when they are compared with each other, even as one grants that Y can have one or neither of the values, K and M—but not *both*—when all three outcomes are considered at once. But in that case, it is clear that Y’s value depends not solely on its internal features, and the particular value bases corresponding to those internal features, B_y , but in part on the alternatives with which it is compared. Thus, either the intransitivity of the “equally as good as relation” suggested by Add is an illusion, or it is based on an Essentially Comparative View of ideals, and not on an alternative version of the Internal Aspects View as Persson suggests.

Next, let me discuss a related way of thinking about the kinds of cases that Persson may have in mind. To do this, it will help to consider Diagram One.

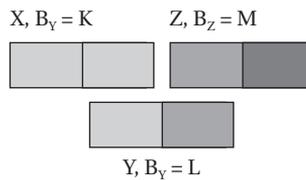


Diagram one

Diagram One represents the values of our three standard outcomes, X, Y, and Z. In accordance with the Internal Aspects View, the value of outcome X is determined by the relevant bases for value that obtain in X, and these bases are a function solely of X’s internal features. However, Diagram One represents a situation where the value of X, which we again represent by the letter K, does not correspond to a single *number*, rather it has a number of varying elements reflecting the ways in which and extent to which X is good. For simplicity, we have assumed that the different aspects of X’s value can be accurately represented by the two connected rectangles of different color and pattern in the top left portion of the diagram. Similarly claims might be made regarding outcomes Y and Z and their values, where the different aspects of Y’s value are represented by the two connected rectangles of different color and pattern in the bottom portion of the diagram, and the different aspects of Z’s value are represented by the two connected rectangles of different color and pattern in the top right portion of the diagram.

As represented, there is considerable overlap in the nature and extent of value between outcomes X and Y, represented by the fact that X’s right rect-

angle, and Y's left rectangle have the exact same color and pattern. Similarly, there is considerable overlap in the nature and extent of value between outcomes Y and Z, represented by the fact that Y's right rectangle, and Z's left rectangle have the exact same color and pattern. But there is no overlap in the nature and extent of value between outcomes X and Z, represented by the fact that X's two rectangles, and Z's two rectangles have completely different colors and patterns.

The following might then be *phenomenologically* accurate. If someone were asked to compare the outcomes X and Y, he might naturally focus on the significant respects in which their values were the same, and thus "perceive" or judge that they had the *same* value. Likewise, if someone were asked to compare the outcomes Y and Z, he might naturally focus on the significant respects in which their values were the same, and thus "perceive" or judge that they had the *same* value. But if someone were asked to compare the outcomes X and Z, he might naturally focus on the fact that their values were not the same at all, and hence "perceive" or judge that they had *different* values. Thus, considering the outcomes two at a time, one might naturally be drawn to judge that X and Y were equally good, and Y and Z were equally good, but that, contrary to the purported transitivity of the "equally as good as relation", X and Z were *not* equally as good.

Here, we might have three outcomes where the judgments people might actually make regarding their value would seem in accordance with principle Add. Moreover, importantly, the values we would be responding to in making our judgments about the different outcomes would be determined solely by the internal features of those outcomes, so it might seem that we can give an account of a violation of the transitivity of the "equally as good as" relation consistent with an Internal Aspects View of ideals.

So should we accept Persson's view, after all? I don't think so; at least not on the basis of the foregoing. My reaction to this kind of case is similar to my reaction to the previous one. My first reaction, and my main one, would be to acknowledge that people might, in fact, react *phenomenologically* to the different outcomes in the way suggested, but to contend that when they did so they were mistaken, and caught in the equivalent of a normative optical illusion. Comparing outcomes X and Y, we might well find the respects in which their values are the same especially *salient*, and this may lead us initially to judge them as equally good. However, once we are clear about what is going on in such cases, it seems clear that X and Y are not *equally* good (or *perfectly* similar regarding value). There are, undoubtedly, respects in which their values are the same, represented by the two rectangles that they each have of *exactly* the same color and shape, but there are also, undoubtedly, respects in which their values are different, represented by the two rectangles that they each have of *completely different* color and shape. Absent a plausible story

that we have not been given for why it is permissible to *completely ignore* the respects in which X and Y differ in the ways and extent to which they are good, it seems hard to stick with the intuitive judgment that X and Y have exactly the *same value, all things considered*. Similar points might be made, of course, about our initial intuitive judgment that Y and Z have the exact same value.

My second reaction to this kind of case would be to point out that, while it would be true that *whatever* aspects of Y's value that we find ourselves responding to arise from Y's internal features, the *particular* features that we focus on in assessing Y's value will *not* solely be a function of Y's internal features. Instead, it will be a function of the alternative outcome with which we compare Y. So, contrary to the Internal Aspects View, we cannot first determine Y's value, considering Y just by itself, do the same for X and Z, and then find out how Y compares with X and Z by comparing them directly in terms of the independent answers we came up with. Rather, in assessing Y's value, we focus and rely on the ways and extent to which Y is good that are represented by the bottom left rectangle in Diagram One when we are comparing Y with X, but we focus and rely on the ways and extent to which Y is good that are represented by the bottom right rectangle of Diagram One when we are comparing Y with Z. Thus, here too, the factors that are relevant and significant for assessing Y's value vary depending on the alternative outcome with which it is compared, and hence it is an Essentially Comparative View, after all, that would account for the plausibility of principle Add and the intransitivity of the "equally as good as" relation, not a rival Internal Aspects View to the version that I presented.

Let me make one final point. In presenting his view, Persson makes it clear that *if* one is going to have an Internal Aspects View that would be compatible with the rejection of the various Axioms of Transitivity, then we have to reject the natural and plausible model for thinking about the goodness of outcomes that I present in my book, where goodness is understood as a property that can, in principle, be quantified and accurately represented by a real number, or a range of real numbers.¹ I agree with Persson about this. Moreover, as I point out in my book, there are various important problems with the "numerical" model in question (see for example Temkin, 2012: ch. 10, note 10). But recognizing this, it is not enough to note that we need something other than my numerical model if we are to explain violations of the Axioms of Transitivity in a way that is compatible with an Internal Aspects View. We need an account of what the alternative way of thinking

1. Parfit has often made similar claims during our discussion about these issues. He eschews thinking about goodness in terms that can be represented by a real number. However, it is not clear to me what the coherent alternative to thinking about goodness in such terms is supposed to be, which fits with the underlying intuitions that motivate the Internal Aspects View in the first place.

about the Internal Aspects View looks like. So far, we don't have even the broadest sketch of such an account —beyond the simple assertion that it can't be like my numerical model.

How, exactly, are we to understand this rival, non-numerical, conception of the Internal Aspects View, so as to capture the various features that Persson and I both agree need to be captured? Here, as elsewhere, the devil is in the details, and I think the burden of proof lies on Persson to further develop and defend the conception he has in mind. Perhaps he thinks he has already done this. But if he has, I am afraid I have missed it. And if I have, perhaps others have as well.

3.

Persson's article presents a striking and intriguing suggestion. He suggests that even on an Internal Aspects View the Axioms of Transitivity should be rejected. Moreover, he contends that this position offers the best way of interpreting what is going on in my book's Spectrum Arguments. Unfortunately, I don't find Persson's claims convincing. As I try to make sense of Persson's view, I keep thinking that either his normative judgments are mistaken —caused, perhaps, by a cognitive illusion— or the *real* explanation for them is provided by an Essentially Comparative View.

In sum, while I welcome further reasons to challenge some of our standard assumptions about practical reasoning, I am not yet persuaded that Persson can deliver on the promissory note that his article offers us. Specifically, I am not yet convinced that there is a plausible rival account of an Internal Aspects View that both fully reflects the position that an outcome's goodness depends *solely* on its internal features, and is compatible with the rejection of the Axioms of Transitivity. Moreover, even if such a view could be defended, I'm not convinced that it would provide the *best* explanation for what is going on in my Spectrum Arguments, rather than the one that I suggested in terms of an Essentially Comparative View. But I look forward to learning more from Persson regarding all of this on another occasion.

BIBLIOGRAPHY

- Broome, J., 1991: *Weighing Goods*, Oxford: Basil Blackwell.
- Persson, I., 2014: "Internal or External Grounds for the Nontransitivity of 'Better/Worse than'", *Law, Ethics and Philosophy* 2: ????
- Rawls, J., 1971: *A Theory of Justice*, Cambridge (Mass.): Harvard University Press.
- Temkin, L., 2012: *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*, Oxford: Oxford University Press.
- 2014: "Rethinking the Good - A Small Taste", *Law, Ethics and Philosophy* 2, this volume: ????

Globalization and global justice in review*

NICOLE HASSOUN

Binghamton University

Abstract

Globalization connects everyone, from the world's poorest slum dweller to the richest billionaire. *Globalization and Global Justice* starts by giving a new argument for the conclusion that coercive international institutions—whose subjects who are likely to face sanctions for violation of their rules— must ensure that everyone they coerce secures basic necessities like food, water, medicines. It then suggests that it is possible for coercive institutions to fulfill their obligations by, for instance, providing international aid and making free trade fair. This overview sketches the argument in the book's first half, as they are the focus of the papers in the symposium.

Keywords: Globalization, global justice, legitimate coercion, autonomy.

1. INTRODUCTION

Globalization connects everyone, from the world's poorest slum dweller to the richest billionaire. US subsidies for ethanol contributed to a world food crisis in 2008 that caused Haiti's government to fall. The subsequent US financial crisis precipitated the European sovereign debt crisis and a global recession felt in even the poorest countries. Today, however, anti-globalization protests pale in comparison to the new protests against economic inequality and oppression that gave rise to the Arab Spring. But many of the new protests, from the US occupy movement to those in Tahrir Square, also focused on what are now—truly global— economic structures' impacts on individuals' ability to meet their basic needs.

The first half of *Globalization and Global Justice (GGJ)* gives a new argument for the conclusion that coercive international institutions—whose subjects who are likely to face sanctions for violation their rules— must en-

* The author would like to thank the editors, commentators, Marcus Arvan, Thom Brooks, and Darrel Moellendorf.

sure that everyone they coerce secures basic necessities like food, water, medicine. Otherwise, the book suggests, these people will not secure sufficient autonomy, which will be defined below to include the basic reasoning and planning capacities necessary to consent, or object, to coercion. The book's second half suggests that it is possible for coercive institutions to fulfill their obligations by, for instance, providing international aid and making free trade fair. It concludes with a new proposal for Fair Trade in pharmaceutical and biotechnology to help people secure access to essential medicines. Because the papers in this symposium focus on the argument in *GGJ*'s first half, however, this overview will do so as well. This *Autonomy Argument* proceeds, roughly, as follows:

- 1) Coercive institutions must be legitimate (i.e. justified in exercising coercive force).
- 2) For a coercive institution to be legitimate it must ensure that its subjects secure sufficient autonomy to autonomously consent to, or dissent from, its rules (henceforth *sufficient autonomy*).
- 3) Everyone, to secure this autonomy, must secure some food and water, and most require some shelter, education, health care, social support, and emotional goods.
- 4) There are many coercive international institutions (that may amount to a coercive international institutional system).
- 5) So, these institutions must ensure that their subjects secure food, water, and whatever else they need for sufficient autonomy.

This argument is intended to address liberals deeply concerned about individual freedom. *GGJ* does not provide an account of individual responsibility for bringing about the requisite institutional change. Nor does it address tradeoffs between fulfilling the condition for legitimacy it defends and other things that matter.¹ Nevertheless, the book attempts to provide reasons for liberals of all sorts—as long as they are deeply concerned about coercion—to accept the *Autonomy Argument*'s conclusion.

Although there are many other good arguments for aiding the global poor, the book attempts to address two kinds of skeptics: Libertarians and statist. Libertarians do not think there are any obligations of global justice or legitimacy to provide aid.² They are, however, deeply concerned about coercion and think no one should have to sacrifice their freedom for others. *GGJ* argues that it is precisely because no one should have to sacrifice their freedom for others that there are significant obligations to the global poor.³ It suggests that if, as many have argued, libertarians should be actual con-

1. I discuss some such tradeoffs elsewhere (Hassoun, 2008) and in this journal.
2. I use the term "libertarian" throughout to refer only to right-libertarians.
3. I discuss elsewhere the relevant sense in which this is true (Hassoun, 2014).

sent theorists, libertarianism entails that people must secure the autonomy they need to consent to coercive rule. So, if the Autonomy Argument goes through, libertarians should agree that coercive institutions must ensure that their subjects secure food, water, and whatever else they need for this autonomy.

Statists often hold that, because states exercise coercion, to be legitimate, they must fulfill significant obligations of global legitimacy or justice. *GGJ* argues that many international institutions also exercise coercion. So, statists should agree that these institutions also have these obligations. What follows recaps the basic line of thought supporting each premise of the Autonomy Argument. In particular, this summary focuses on a line of argument for the second premise, which is most relevant when addressing libertarians. I suspect that this premise is the most controversial—and potentially important—aspect of the argument. In the book, I distinguish between the defense of the argument by appeal to libertarian principles (the topic of Chapter 3) and the more general defense (discussed in Chapter 2) by renaming the Autonomy Argument the *Legitimacy Argument*. Here, however, I will not make much of the distinction and simply highlight some of the argumentative moves intended to address libertarians.

2. THE FIRST PREMISE

GGJ defines institutional legitimacy in this way: A coercive institution is legitimate only if it has the right to use coercive force.⁴ *Legitimacy*, then, is a “justification” right to rule through force (Landenson, 1980). Having a justification right is having a moral permission to make coercive rules and give coercive commands. Legitimacy, in this sense, must be distinguished from justified authority (Christiano, 2004). A coercive institution has justified authority if and only if it is legitimate and individuals have a moral duty to comply with its rules (Buchanan, 2004: 237). Some rights may carry with them correlative duties (Simmons, 1979). *GGJ*’s argument does not rely on it being the case, however, that whenever coercive institutions have a right to rule through force, their subjects are obligated to obey their dictates (though this may be so).

4. Legitimacy, as the book uses the term, comes in degrees. Some people believe legitimacy is an all or none affair. This is not a substantive disagreement. Those who hold a binary theory of legitimacy can specify that an institution is legitimate in the binary sense if it surpasses a threshold of legitimacy in the degree sense. However, the degree conception allows for different thresholds to be specified for different purposes. For the purpose of the Autonomy Argument, one need only suppose that imperfectly legitimate institutions must be reformed. I owe thanks to Allen Buchanan for discussion on this point.

According to the first premise of the Autonomy Argument, any coercive institution must be legitimate. Although this point is relatively uncontroversial, here is an argument in its defense that is intended to appeal to liberals deeply concerned about coercion. Following John Locke, one may hold that each person has a natural right to freedom and so, rights-respecting people cannot be subject to others' commands without justification (Locke, 1690). H.L.A. Hart provides one way of defending a natural right to freedom. Hart argues that if there are any natural rights, there is a natural right to freedom (Hart 1955). Alternately, one might try to ground the concern for freedom in a concern for individuals' interests or autonomy. But, since the Autonomy Argument is not intended to address skeptics about the importance of freedom, *GGJ* does not examine the alternatives at great length.

3. THE SECOND PREMISE

According to the second premise of the Autonomy Argument, to be legitimate, coercive institutions must ensure subjects secure sufficient autonomy. Consider what this means.

First, people are subject to a coercive institution when the rules of the institution apply to them and to secure *sufficient autonomy*, people must be able to reason about, make, and carry out some significant plans on the basis of one's desires, beliefs, values, and goals (henceforth *commitments*). More precisely, people must be able to reason about, make, and carry out the plans necessary to consent or object to the coercive institutions to which they are subject. To secure sufficient autonomy people need not be perfectly autonomous. People need only possess a few *conditions* for autonomy. The book appeals to these conditions for autonomy to secure broad agreement on the Autonomy Argument's second premise. At least it is not plausible to reject this premise because the conditions for autonomy it relies upon are too demanding. Those who accept fuller (e.g. Kantian) conceptions of autonomy might run a similar argument for more significant obligations to the global poor.

Second, what is necessary to *ensure* that people secure sufficient autonomy will vary with the case. Coercive institutions must do whatever is necessary (and permissible), to ensure that their subjects *become and remain autonomous until and unless they autonomously relinquish their ability to do so*. What is necessary depends on how close one is to being able to secure such autonomy and what resources one already has. In cold climates, for instance, one may need to secure heat. In the tropics, heat is usually unnecessary. Some will be able to secure sufficient autonomy as long as they are free from interference. Others, however, need assistance to secure sufficient

autonomy. The coercive institutions to which these people are subject may have to provide this assistance. If, for instance, a person is in a coma from which she could recover with proper medical care and she is not receiving such care from friends, family, or benefactors, then the coercive institutions to which she is subject must provide it. These institutions have a responsibility of last resort. An institution does not lose legitimacy if it does not help someone secure sufficient autonomy and this person does not have the potential to secure such autonomy. When they are very young, children are not able to secure any autonomy at all. Most children who receive proper care will be able to secure sufficient autonomy as they get older. If no one else does so, a legitimate coercive institution must help these children secure such autonomy once they are old enough.

It is possible to defend the condition for legitimacy in the Autonomy Argument's second premise in two steps. What follows will argue, first, that coercive institutions can only be legitimate if as many of their subjects as possible secure sufficient autonomy. Second, it will argue that such institutions must do what they can to ensure subjects secure this much autonomy. It is possible to defend the first claim by appeal to the nature of liberalism (leaving the possibility constraint implicit where its importance is minor). The second claim follows from the first and some observations about the nature of ensuring and coercive institutions.

3.1. The First Point Necessary for Establishing the Enabling Condition

At the heart of liberalism is the concern for individual freedom. Recently liberals have focused primarily on arguing that coercive institutions must be decent, if not fully just (Rawls, 1993; Pogge, 1989). An equally powerful strand in liberal thought, however, expresses the idea that the *actual* relationship between the rulers and *each ruled person* must be voluntary in some way.

Liberals deeply concerned about individual freedom disagree about what makes the relationship between the rulers and the ruled voluntary. They all agree, however, that this relationship can only be voluntary if the ruled possess at least some freedom. This freedom is not constituted by the social order and it is compatible with significant constraints on social life (Waldron, 1987: 133). The key idea is that subjects must be free to determine their actions and shape the nature of their relationship with the coercive institutions to which they are subject (Waldron, 1987: 132). Although individuals may not have a choice of whether or not they are subject to a coercive institution, subjects must be able to exercise some control over the way they react to their subjection. Subjects should get to decide whether or not to abide by, dissent from, or consent to coercive institution for themselves (Waldron,

1987: 146). Political liberals almost unanimously agree, for instance, that people have a right to dissent from the rule of the coercive institutions to which they are subject by conscientious objection, non-violent protest, passive resistance, and so forth.

To consent to, or dissent from, coercive institutions in these ways, people must be able to reason about, make, and carry out some significant plans in light of their beliefs, desires, values, and goals; they must be able to secure sufficient autonomy.⁵ (Recall that the conditions for sufficient autonomy are quite minimal - people secure sufficient autonomy when they secure basic reasoning and planning capacities). So liberals implicitly accept the first claim embodied in the enabling condition for legitimacy; those living under coercive institutions must be able to secure sufficient autonomy for the coercive institutions to which they are subject to be legitimate. William Kymlicka puts the point this way: “liberalism is committed to (and perhaps even defined by) the view that individuals should have the freedom and capacity to question and possibly revise the traditional practices of their community, should they come to see them as no longer worthy of their allegiance” (Kymlicka, 1992). The book explains, at some length, why liberals of many persuasions should accept the first part of the second premise of the Autonomy Argument; for coercive institutions to be legitimate, their subjects must secure sufficient autonomy.

Consider, here, just why libertarians, in particular, should endorse the first part of the second premise of the Autonomy Argument. There is a well-known argument in the literature for the conclusion that libertarians should be actual consent theorists. Very roughly, on the relevant version of actual consent theory, coercive institutions must, insofar as possible, secure their rights-respecting subjects’ consent until, and unless, they give up the right to consent. Any agent, or institution, may be justified in coercing those who violate others’ rights. To use an example from John Simmons, even “the Third Reich was justified in prohibiting rape and punishing rapists” (Simmons, 1999). But coercive institutions usually do more than this. When they create norms, rules, and procedures governing the use of force, for instance, they prevent people from defending their own rights. This is clearly the case for (even libertarian) states, which claim a monopoly on coercive force within a territory traditionally defined. Since libertarians hold that everyone has a basic right to defend their rights, consent is required for such coercion. Assuming this argument goes through, *GGJ* notes that, in order to actually consent, people must be able to do so. This requires at least basic reasoning

5. Recall that this just presupposes some minimal conditions for full autonomy—one need not have coherently structured values e.g. to have the basic reasoning and planning capacities at issue.

and planning capacities sufficient to autonomously consent (i.e. sufficient autonomy).

3.2. The Second Point Necessary for Establishing the Enabling Condition

Why must coercive institutions do what they can to ensure subjects sufficient autonomy? The preceding argument entails that when coercive institutions subject people who cannot secure sufficient autonomy to coercive rules and do not do what they can to ensure subjects secure this autonomy, they are illegitimate. This is because coercive institutions are not justified in exercising rights—constraining coercive force over rights—respecting people who could, but have not, secured sufficient autonomy. Yet coercive institutions exercise such force. If coercive institutions continue to exercise coercive force, legitimacy requires that they do what they can to ensure subjects secure sufficient autonomy. Coercive institutions do continue to exercise this force (insofar as they remain coercive institutions). So, they must do what they can to ensure subjects secure sufficient autonomy.

There are a few caveats to this conclusion. Others may have primary responsibility for enabling those subject to coercive institutions to secure sufficient autonomy. Moreover, if people secure this autonomy on their own, or with the help of friends and/or benefactors or give up their right to do so, the coercive institutions to which they are subject need not do a thing. These coercive institutions must generally step into the breach, however, if help is required. It is only if they do this that as many of their subjects as possible will secure sufficient autonomy. There may also be other conditions for institutional legitimacy. Still, coercive institutions may be justified in doing other things before enabling their subjects to secure sufficient autonomy. This is especially likely if enabling people to secure such autonomy is extremely costly.

4. THE THIRD PREMISE

The final premise of the Autonomy Argument is this: Most people must at least be able to secure some minimal amount of food, water, shelter, education, health care, social and emotional goods to secure sufficient autonomy. Recall that, to secure the sort of autonomy at issue, people must at least be able to reason about, make, and carry out some significant plans on the basis of their commitments.

Even without explaining these conditions for autonomy in any detail, it should be clear that those who lack basic food, water, shelter, education and health care are likely to suffer from autonomy undermining disabili-

ties. Malnutrition inhibits one's immune system's ability to fight infection and poor nutrition is linked to many non-infectious illnesses.⁶ Similarly, if people lack adequate shelter, they may be exposed to environmental hazards including disasters, pollutants, parasites, and bacteria, like dysentery, tetanus, typhoid, cholera, or hepatitis from in flood water or unsanitary living conditions (Red Cross, 2007). Those without basic health care, food or water are at risk of diseases causing disabilities or premature death (*ibid.*) incompatible with securing sufficient autonomy.

Less obviously, those without basic education, emotional and social goods may suffer from autonomy undermining disabilities (Woolcock, 2001; Doyle, 2002). Basic education, emotional and social goods are often necessary for securing decent living conditions, health care, livelihood opportunities, and earning power (Marmot, 2004). Those who lack (formal or informal) elementary education may not develop, or maintain, the reasoning and planning skills they need to secure sufficient autonomy.⁷ Those who lack basic emotional and social goods, like self-esteem, are at high risk for mental and physical illness, suicide, and early death from other causes (Cullen and Whiteford, 2001; Brock, 1999; Hudson, 2005; Woolcock, 2001). "Fear, insecurity, dependency, depression, anxiety, intranquility, shame, hopelessness, isolation and powerlessness... such experiential elements of a bad life... [often impact] ... agency." It is true that some people are able to secure sufficient autonomy without being able to obtain even minimal education or social or emotional goods. But, this kind of severe deprivation will undermine most people's ability to secure sufficient autonomy.

5. THE FOURTH PREMISE

Before arguing that there are many coercive international institutions, *GGJ* sketches a conception of coercion. It explains that an institution is *coercive* (see Gerald 2003) when individuals, or groups, violating its dictates are likely to face sanctions for the violation. A *sanction* is a punishment or penalty.

6. Scurvy results from a lack of vitamin C, beri-beri from a lack of thiamine, pellagra from niacin deficiency, and macrocytic and microcytic anemia from folic acid and iron deficiencies, for instance. There is also a lot of evidence that decent nourishment is important for good cognitive functioning. Children's mental functioning can even be impaired if their mothers do not receive proper nourishment during pregnancy. Keratomalacia which results from vitamin A deficiency, kwashiorkor which results from protein deficiency, and iodine deficiencies can all lead to severe disabilities and death. See Leathers and Foster, 2004.

7. Stress may contribute to a host of autonomy-undermining mental disorders. Stress can, for instance, cause panic attacks and depression. Psychological disorders can reduce the ability of one's immune system to fight infection. See Beaton, 2003. The causal evidence suggests that perception of low social standing may increase stress which reduces immune functioning and can harm health in other ways as well.

Coercion usually creates conditions under which the coerced have no good alternative except to do what their coercer wants them to do. This is usually explained by the fact that the coerced are threatened by sanctions (Risse, 2006). Depending on the kind and amount of coercion and so forth, coercion may or may not undermine autonomy to any significant degree. Usually, it engages the will of the coerced. Still, people can be coerced into doing what they would otherwise do freely. Furthermore, institutions can be coercive even if they do not coerce anyone into doing anything. If, for instance, a state only creates just laws and everyone willingly obeys, it may still be coercive. The state is subjecting people to coercive laws, though it never has to sanction anyone for disobedience.

GGJ notes that the preceding analysis leaves a lot open. A lot hangs on what counts as a violation, a punishment or penalty, and a good alternative (Anderson 2006). Some hold that only threats can be coercive while others say sanctions can include withholding an offered good. There is also disagreement about the appropriate baseline relative to which something counts as a sanction. It is not clear, for instance, whether one can be sanctioned in ways that do not violate rights.

The book allows that some international institutions are not coercive. Non-binding treaties like the *Declaration on the Rights of Disabled Persons* are not coercive. Nor are non-governmental organizations that offer only voluntary programs, normally, coercive.

GGJ argues, however, that there are many coercive international institutions. It does so by providing examples that should appeal to those with widely divergent accounts of coercion. Many international institutions' dictates are binding and non-voluntary.

(Since libertarians tend to think there is a lot of coercion in international affairs, what follows will not focus on addressing libertarians, in particular, though the book provides additional examples that are intended primarily to address libertarians.)

There are many ways the international institutions governing trade exercise indirect coercion. Institutions like the World Trade Organization (WTO) and the North American Free Trade Agreement (NAFTA) impose sanctions on countries that violate property rights or the rules of the market. States enforce these sanctions. The NAFTA sanctioned Mexico for prohibiting Metalclad from operating a toxic waste dump in San Luis Potosi, for instance. Mexico had to pay Metalclad 16 million US dollars in damages (Wallach, 2005). The WTO found the US guilty of violating its rules with the Byrd amendment. It allowed prosecuting countries to impose import duties on the US until the US repealed the act (European Union, 2005). Recently, the WTO sanctioned the European Union (EU) by allowing the US to impose

tariffs on EU goods because the EU had used import licensing requirements to support Caribbean banana producers (British Broadcasting Company, 1999). In many cases, laws passed by states as a result of WTO rulings eventually coerce businesses and individuals into abiding by the rulings.

The United Nations (UN) also exercises indirect coercion. The UN Security Council imposes economic sanctions, air traffic controls, and arms embargos on countries, and groups within countries, that threaten international security. The UN has, for instance, sanctioned Rhodesia, Iraq, South Africa, Serbia, Montenegro, Yugoslavia, Somalia, Libya, Haiti, Sudan, Rwanda, Sierra Leon, Ethiopia, Eritrea, and groups within Cambodia, Angola, and Afghanistan (Roberts, 2001). The UN Security Council also authorizes the use of force against countries threatening international peace. When Iraq invaded Kuwait the UN authorized the use of force to stop the invasion. The UN-mandated International Security Assistance Force of about 30,000 troops is currently involved in military action in Afghanistan (United Nations, 2003).

Moreover, many countries' participation in international institutions is not voluntary. Countries often pay significant penalties if they do not abide by WTO, UN, World Bank or International Monetary Fund (IMF) rules. Sometimes countries do not have other good options and so are not free to resist these organization's conditions. Highly indebted poor countries facing default, for instance, may have to abide by IMF conditionality. Furthermore, some argue that international institutions bear responsibility for poor countries' having no reasonable option but to abide by their rules by having contributed to their impoverishment.⁸ At least in such cases, international institutions are indirectly coercing individuals in the way that a man with a gun indirectly coerces someone if he forces another person to threaten the first.⁹

Many international institutions also exercise direct coercion. UN peacekeeping forces exercise direct coercion by, for instance, taking over territory, patrolling borders, and creating safe havens for refugees. Those who attempt to wrest control from the UN, or enter its protectorates or safe zones without permission, face sanctions for the violation. Peacekeeping forces have been deployed in places as diverse as Congo, Iran, Lebanon, Sinai, Yemen, the Golan Heights and Cyprus. Between 1988 and 1999 alone, the UN initiated forty peacekeeping missions (Roberts, 2001).

8. This condition may not be necessary for coercion. For discussion see Nozick, 1969; Zimmerman, 1981; Gorr, 1986; McGregor, 1998/89; Held, 1972; Van De Veer, 1979.

9. Even if individuals' states have other options and are thus partly responsible for coercing their people, international institutions may still be acting wrongly. Knowing how states are likely to act, it may not be acceptable for these institutions to act in the ways that they do. But this paper sets this point aside.

Other international institutions also coerce individuals directly, sometimes in ways that violate rights. Consider, for instance, what happened as the humanitarian crisis in the Balkans developed. The UN imposed an arms embargo against the former Yugoslavia, a flight ban over Bosnia and Herzegovina, and economic sanctions against Montenegro and Serbia. NATO enforced these measures. In 1999, when the UN peacekeeping force failed to prevent the Srebrenica massacre, NATO bombed Bosnia. NATO then enforced the Bosnia-Herzegovina peace agreement under the auspices of a UN protectorate and brought individuals accused of war crimes to The Hague. It thereby directly coerced, and enabled the court to coerce, individuals. Eventually, NATO ceded command in Bosnia to the EU, which deployed its own troops (NATO, 2007).

6. CONCLUSION

Many people resist the idea that there are any obligations of justice to the global poor. If the Autonomy Argument goes through, these people are mistaken. Legitimacy requires that coercive institutions do what they can to ensure that all of their subjects with the potential to secure sufficient autonomy secure adequate food, water, shelter, education, health care, social and emotional goods. The second half of *GGJ* argues that there are many things we can do to help people secure what they need for sufficient autonomy. So, the book concludes, in a world where 18 million people die annually of easily preventable poverty-related causes, there are many things we can and must do now to ensure that everyone secures what they need.¹⁰

BIBLIOGRAPHY

- Anderson, S., 2006: "Coercion", *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/coercion/>.
- British Broadcasting Company, 1999: "WTO Approves Banana Sanctions", <http://news.bbc.co.uk/2/hi/business/322938.stm>.
- Brock, K., 1999: "Its Not Only Wealth that Matters it's Peace of Mind Too': Review of Participatory Work on Poverty and Illbeing", Institute of Development Studies: Birmingham.
- Buchanan, A., 2004: *Justice, Legitimacy, and Self-determination: Moral Foundations for International Law*, Oxford: Oxford University Press.

10. See World Health Organization 2004. Although the Autonomy Argument does not establish this conclusion, coercive institutions must also allow individuals to meet their basic needs in a decent, legitimate way. No one should have to scavenge under burning heaps of garbage to survive.

- Christiano, T., 2004: "Political Authority", *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/authority/>.
- Cullen, M., and Whiteford, H., 2001: "Inter-relations of Social Capital with Health and Mental Health", Mental Health and Special Programs Branch Commonwealth Department of Health and Aged Care Discussion Paper. Canberra: Commonwealth Department of Health and Aged Care.
- Doyle, R., 2002: "Calculus of Happiness: Assessing Subjective Well-being Across Societies", *Scientific American* 287/5 :3.
- European Union, 2005: "U.S. Congress Repeals Byrd Amendment but Allows for a Transition Period", No. 128/05.
- Gaus, G., 2003: "Liberal Neutrality: A Compelling and Radical Principle", in *Perfectionism and Neutrality: Essays in Liberal Theory*, S. Wall and G. Klosko (eds.), New York: Rowman and Littlefield Publishers.
- Gorr, M., 1986: "Toward a Theory of Coercion", *Canadian Journal of Philosophy* 16, 3: 383-406.
- Hart, H. L. A., 1955: "Are There Any Natural Rights?", *The Philosophical Review* 64: 175-191.
- Hassoun, N., 2009: "Meeting Need", *Utilitas* 21, 3: 250-275.
- 2014: "Coercion, Legitimacy, and Individual Freedom: A Reply to Soderholm", *Journal of Philosophical Research* 39: 191-198.
- Held, V., 1972: "Coercion and Coercive Offers", *Coercion: Nomos* 14.
- Hudson, C. G., 2005: "Socioeconomic Status and Mental Illness: Tests of the Social Causation and Selection Hypotheses", *American Journal of Orthopsychiatry* 75, 1: 3-18.
- Kymlicka, W., 1992: "The Rights of Minority Cultures: Reply to Kukathas", *Political Theory* 20, 1: 142.
- Landenson, R., 1980: "In Defense of a Hobbesian Conception of Law", *Philosophy and Public Affairs* 9, 2: 134-159.
- Leathers, H., and Foster, P., 2004: *The World Food Problem: Tackling the Causes of Undernutrition in the Third World*, Colorado: Lynne Rienner Publisher.
- Locke, J., 1690: *Second Treatise on Civil Government*, 1990, ed. C. B. MacPherson, Indianapolis: Hackett.
- Marmot, M., 2004: *Status Syndrome: How your Social Standing Directly Affects your Health and Life Expectancy*, London: Bloomsbury.
- McGregor, J., 1998/1989: "Bargaining Advantages and Coercion in the Market", *Philosophy Research Archives* 14: 23-50.
- NATO 2007: "What's on NATO's Agenda?", North Atlantic Treaty Organization.
- Nozick, R., 1969: "Coercion", in *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, ed. W. Morgenbesser, St Martin's Press, 440-472.
- Pogge, T., 1989: *Realizing Rawls*, New York: Cornell University Press.
- Rawls, J., 1993: *Political Liberalism*, New York: Columbia University Press.
- Red Cross, 2007: "American Red Cross Urges Public Health Precautions", Red Cross: Washington D.C.
- Risse, M., 2006: "What to Say about the State", KSG Working Paper No. RWP06-008, Cambridge, Mass.: Harvard University.
- Roberts, A., 2001: "United Nations", in *The Oxford Companion to Politics of the World* (Second Edition), ed. J. Krieger, Oxford: Oxford University Press.

- Simmons, J., 1979: *Moral Principles and Political Obligations*, Princeton: Princeton University Press.
- 1999: “Justification and Legitimacy”, *Ethics* 109, 4: 739-771.
- United Nations, 2003: “Security Council Seeks Expansion of Role of International Effort in Afghanistan, to Extend Beyond Kabul”, Press Release SC/7894. Security Council 4840th Meeting (PM) 13/10/2003. <http://www.un.org/News/Press/docs/2003/sc7894.doc.htm>.
- Van De Veer, D., 1979: “Coercion, Seduction, and Rights”, *The Personalist* 58: 374-381.
- Waldron, J., 1987: “Theoretical Foundations of Liberalism”, *Philosophical Quarterly* 37, 147: 133.
- Wallach, L., 2005: “Slow Motion Coup d’Etat: Global Trade Agreements and the Displacement of Democracy”, *Multinational Monitor* 26: 1-2, <http://multinational-monitor.org/mm2005/012005/wallach.html>.
- Woolcock, M., 2001: “The Place of Social Capital in Understanding Social and Economic Outcomes”, *Isuma* 2, 1, <http://www.oecd.org/innovation/research/1824913.pdf>.
- World Health Organization, 2004: “WHO ‘Preparing for Treatment’ Programme”, Call for Tenders to WHO, Annex Table 2, Geneva: World Health Organization.
- Zimmerman, D., 1981: “Coercive Wage Offers”, *Philosophy & Public Affairs* 10, 2: 121-145.

Libertarian Welfare Rights: Can We Expel Them?

CHARLES GOODMAN

Binghamton University

Abstract

In *Globalization and Global Justice*, Nicole Hassoun presents a new and fundamental challenge to libertarian political thought. Her Legitimacy Argument tries to show that natural rights libertarians are committed by their own principles to a requirement that their states recognize and meet the positive welfare rights of certain merely potentially autonomous persons. Unfortunately, this argument suffers from two flaws. Hassoun needs to show, but has not shown, that the libertarian state would have to infringe any of the negative rights of the merely potentially autonomous in such a way as to require consent from them. Moreover, the libertarians could arrange their institutions, justifiably by their own lights, so as to expel all indigent, merely potentially autonomous persons from their territory. This second solution is intuitively unpalatable, but may be no more morally problematic than the basic natural rights libertarian view itself.

Keywords: libertarianism, positive rights, Nicole Hassoun, autonomy, John Locke.

1. INTRODUCTION: THE LEGITIMACY ARGUMENT

Much recent work on global justice has focused on attempts to convince libertarians that the processes of globalization generate significant obligations to help the global poor. The work of Thomas Pogge and others towards this goal has recently been supplemented by Nicole Hassoun's important book *Globalization and Global Justice*. Hassoun presents a new, fundamental, and apparently devastating challenge to libertarian political thought. According to Hassoun's Legitimacy Argument, natural rights libertarians who reject anarchism, defend the state's monopoly on force, and accept actual consent theory, cannot explain how it is morally legitimate for them to coerce certain non-autonomous but potentially autonomous persons. In order for a libertarian state to exercise jurisdiction over these persons, it must provide them, insofar as it can, with what they need to become autonomous. Thus the nor-

mative legitimacy of even a libertarian state would depend on its practical recognition of certain positive welfare rights.

Hassoun focuses our attention, as other writers have, on the horrifying moral tragedy of our time: the plight of the global poor in a world that contains so much affluence. Libertarians are aware of this tragedy, and have no need to minimize it. They propose to relieve the misery of the poor by extending free trade, the rule of law, and the institutions of capitalist society to every part of the globe. Since these institutions have repeatedly succeeded where no others have in transforming poor countries into rich countries, they constitute the only approach on which we have any reason to rely for saving the world's poor from their wretched condition.

Hassoun would disagree with this program; and she offers various empirical arguments for the importance of foreign aid in helping to alleviate global poverty. But regardless of the outcome of the empirical debate about what measures would most effectively help the poor, libertarians must still contend with Hassoun's case for the claim that, contrary to their view, potentially autonomous people have positive welfare rights that can be grounded in considerations about individual freedom and consent that libertarians themselves accept, and that governments cannot be legitimate if they do not appropriately respond to these rights.

Libertarians advocate the creation of minimal states that do nothing other than protect the negative rights of their citizens. Such states would have police forces, courts, and armies, and would use them to punish crime, deter aggression and enforce contracts, but would not collect taxes from citizens for any other purpose beyond these. Call a minimal political institution of this type a *libertarian state*. Those who claim that only a libertarian state would be legitimate, because any more extensive state would violate the natural rights of citizens, I will call *natural rights libertarians* (see Nozick, 1977.)

Two prominent theorists, Thomas Pogge and James Sterba, have attempted to show that the natural rights libertarian position, as just explained, is unstable: under contemporary conditions, their view should imply stringent, enforceable duties to help the global poor. For example, Pogge draws on the Lockean understanding of property rights to try to show that unless we provide the global poor with considerably more resources than they now possess, we will be violating their negative rights (Pogge, 2002: 208-9.) Sterba claims that by enforcing property rights, the libertarian state violates the poor's right to take what they need to survive; this conflict of rights should be resolved in favor of a duty to aid (Sterba, 2005: 47-48.) But, as Hassoun points out, libertarians have been unconvinced by these arguments. Some "reject Sterba's conclusion because they do not believe that a conflict

of rights generates a duty to aid the poor.” Others have tried to show “that libertarians are likely to reject Pogge’s baseline for harm.” (Hassoun, 2012: 91.) Thus, the arguments of Pogge and Sterba remain controversial. Hassoun therefore proposes her own, new argument, which attempts to show in an entirely different way that the basic moral principles that underlie natural rights libertarianism should, properly understood, require institutions that redistribute resources towards the poor. Although this argument is framed in ways that address the global poor, the core of the argument is applicable to a single libertarian state. Thus, despite the fact that Hassoun’s book as a whole is directed to addressing issues of global justice, I will be focusing on her argument primarily as it applies domestically.

On p. 92 of *Globalization and Global Justice*, Hassoun states her Legitimacy Argument as follows:¹

- (1) Coercive institutions must be legitimate.
- (2) Roughly, for a coercive institution to be legitimate it must ensure that its subjects secure sufficient autonomy to autonomously consent to, or dissent from, its rules (henceforth *sufficient autonomy*.)
- (3) Everyone, to secure this autonomy, must secure some food and water, and most require some shelter, education, health care, social support, and emotional goods.
- (4) There are many coercive international institutions.
- (C) So, these institutions must (roughly) ensure that their subjects secure food, water, and whatever else they need for autonomy.

I will not be questioning any of premises 1, 3, or 4. Premise 1 is accepted, in some form, by nearly everyone. Premises 3 and 4 look like straightforward empirical truths. So what supports premise 2?²

According to Hassoun, natural rights libertarians should accept actual consent theory: they should hold that coercive institutions are legitimate only if those subject to them have actually consented to their rule. She persuasively argues that the standard objections against actual consent theory should not be acceptable to libertarians. So a libertarian minimal state must secure the actual consent of autonomous persons that live on its territory.

1. The “Legitimacy Argument”, as discussed here, is a form of the more general Autonomy Argument that has been adapted to apply to libertarian theories. Although Hassoun believes, and attempts to show in her Chapter 2, that all persons, whether autonomous or not, do have positive welfare rights, the intention of the Legitimacy Argument is to show specifically that, even on purely libertarian assumptions, persons who are merely potentially autonomous would have positive welfare rights against a libertarian state.

2. Premise 2 starts with the word “roughly.” Hassoun explains the qualifications to this premise that she thinks are necessary at 93-94. They involve those who do not respect the rights of others; those who can never become autonomous; and those who somehow give up their right to consent. None of these qualifications will be relevant to the argument of this paper.

But what about those who are merely potentially autonomous, such as children, or the curably mentally ill? They are still subject to the coercive force of the law of a libertarian state. In order for the state to be justified in infringing their rights through coercion, Hassoun argues, it must do what is necessary to get their actual consent. But since these merely potentially autonomous individuals do not presently have the normative capacity to grant valid consent, so long as they remain as they are, the state cannot get what it needs from them. So it is morally required to provide them with whatever they need in order to become autonomous, so that they can eventually consent to its rule. This Legitimacy Argument would, then, require that the curably mentally ill should receive treatment at public expense, and that children should have their basic needs met and should be provided publicly with sufficient education to become autonomous. Since these goods and services will be paid for from taxation, the Legitimacy Argument entails that citizens of a minimal libertarian state have robust positive duties towards the merely potentially autonomous; should these duties not be fulfilled, their state becomes normatively illegitimate.

To derive premise 2 from actual consent theory, Hassoun needs the plausible assumption that even a libertarian state would have to employ coercion against those living in its territory, including those who are only potentially autonomous. But how, exactly, would the libertarian state use coercion against merely potentially autonomous citizens? When we separately examine the various rights that non-autonomous persons might have, we can identify what may be a serious flaw in this argument.

2. FIRST REPLY: THE RIGHTS OF THE MERELY POTENTIALLY AUTONOMOUS

For present purposes, we can say that, from the kind of libertarian perspective we are examining, individuals have natural rights to bodily integrity, property, self-defense, and punishment. The right to bodily integrity is a trivial consequence of the basic libertarian premise of self-ownership. Rights to property are the result of the appropriation of unowned natural objects and their transformation through labor. Various widely accepted and uncontroversial human rights, such as rights to free speech, free association, due process in criminal cases, and so on, are seen by libertarians as flowing from these two more basic rights to bodily integrity and property. The other two basic natural rights authorize responses to rights violations by others. The right to self-defense gives us limited permissions to respond to violations that are occurring in the present or are likely to occur in the future; the right to punish gives us limited permissions to respond to violations that have occurred in the past.

There is no clear reason why a libertarian state would have to commit aggression against non-rights-violating potentially autonomous persons in its territory, depriving them of their rights to bodily integrity. Nor would the libertarian state arbitrarily confiscate their property. But there is a problem about whether it would be permitted to ask the merely potentially autonomous to pay for its protective services, since they are unable to give valid consent to the contract that authorizes payment for such services. Of course, the citizens of such a state could agree to offer protective services to the potentially autonomous for free. In the case of merely potentially autonomous people who have no valuable resources, this may be the only option. But for those merely potentially autonomous citizens who own some resources, perhaps through inheritance or gift, another approach may be available.

Since they choose to pay the fee for protection, the autonomous citizens of the libertarian state apparently consider that the benefits they receive from this protection outweigh the costs. This is not because of idiosyncratic preferences on their part. The human need for protection against violence is as widely shared as the needs for shelter and food. Thus, we can safely be confident that, if potentially autonomous people receive protection from the state at the expense of being required to pay for it, they will benefit, on net, from the exchange. So perhaps the state can coerce them to pay taxes, or require their guardians to pay taxes on their behalf and out of their property, for paternalistic reasons. Obviously libertarians protest vigorously against paternalist coercion directed against rational adults; but it should be almost equally obvious that libertarians are not required by the logic of their position to reject the paternalist coercion of small children, animals, or other non-autonomous or merely potentially autonomous beings with interests. Hassoun restricts her argument to libertarian views that do not make it easy to justify coercion of the potentially autonomous for the benefit of others (90,) but that restriction does not rule out views that allow coercing such persons for their own benefit. I conclude from these considerations that, in requiring those potentially autonomous beings on their territory who own valuable resources to contribute some of those resources to the defense of their society from violence, the libertarians would not wrong those beings.

Would the libertarian state need to take away the potentially autonomous inhabitants' right to self-defense? It's not clear why it would. Most likely, citizens of a libertarian state who had the appropriate cognitive and physical abilities would retain a robust right to defend themselves against actually occurring violent attacks, especially in emergencies when the police or other agents of the state happen not to be present. Some individuals who lacked the capacity for full autonomy might, in spite of their immaturity or cognitive impairments, still be capable of accurately perceiving that they were under violent attack and of defending themselves in a proportionate man-

ner. The libertarian state could recognize the right to self-defense of both its autonomous and non-autonomous citizens in the same way. Of course there might be some inhabitants of the state's territory who are so young, or so gravely mentally impaired, that they are unable to recognize when they are under attack, or to assess the seriousness of the attack so as to be able to respond in a reliably proportionate way. But, libertarians could argue, those non-autonomous individuals don't have a right to self-defense, and so the libertarian state can't be accused of depriving them of any right by forbidding them to defend themselves.

That leaves only the right to punish as a potential source of Hassoun's moral criticism. The libertarian state would insist on requiring all inhabitants of its territory to surrender the individual right to punish that, according to natural rights theory, they would have had in the state of nature, and providing them instead with a claim to impartial justice as administered by its courts. It's primarily for this reason that the natural rights libertarian position is not a form of anarchism. But why is it permissible to take away, without consent, the right to punish held by the merely potentially autonomous? This deprivation can't be justified on paternalist grounds. When the libertarian state requires a specific individual to surrender her right to punish, the motivation for doing so is not primarily to benefit that individual herself, but rather to protect others from the consequences of being punished by her in a biased or disproportionate manner. So if the merely potentially autonomous have a right to punish at all, then depriving them of it poses a serious moral problem. But do they?

Locke would have said, I think, that they do not. To have a right to punish in a Lockean state of nature, a being must be capable of knowing the law of nature and of administering punishment in a proportionate manner. Thus Locke writes that a man in the state of nature has

no absolute or arbitrary power, to use a criminal, when he has got him in his hands, according to the passionate heats, or boundless extravagancy of his own will; but only to retribute to him, so far as calm reason and conscience dictates, what is proportionate to his transgression, which is so much as may serve for *reparation* and *restraint* (Locke, 1982: II.8. Emphasis in original.)

To exercise this limited power of punishment, someone would seem to need those faculties which Locke calls "calm reason and conscience." Indeed, the proper use of punishment by a being in the state of nature would seem to require a fairly sophisticated deployment of rationality. Locke writes that "each transgression may be *punished* to that *degree*, and with so much *severity* as will suffice to make it an ill bargain to the offender, give him cause to repent,

and terrify others from doing the like” (Locke, 1982: II.2.) Some might question whether an adult of ordinary intelligence, or even a panel of experts on criminology, would be able to do a good job of ascertaining the correct punishment for a particular offense, on this conception. But Locke, defending his conception of the law of nature as including a scale of appropriate punishments, insists that “it is certain there is such a law, and that too, as intelligible and plain to a rational creature, and a studier of that law, as the positive laws of commonwealths, nay possibly plainer...” (Locke, 1982: II.12.) However plausible or implausible this may be, Locke surely could not have claimed, without recourse to innate ideas or other views that he explicitly denied, that a human who has not yet become rational would nevertheless be able to know this law. And if you can’t know, even approximately, what punishments it would be appropriate to administer, you can’t have or claim a right to punish others, even if they encroach upon your undoubted rights.³

When the right to punish is understood in this Lockean way, it is unclear that there would ever be a case of someone who was merely potentially autonomous, in the sense of being insufficiently rational to have the normative power to consent to the authority of the state, who would yet be sufficiently rational in certain specific respects so as to have a right to punish. Nevertheless, suppose that, in rare cases, this condition is met. Someone meeting this condition who was deprived of the right to punish by a libertarian state would be no more wronged than an unusually mature and responsible fifteen-year-old who is denied the right to drive a car in today’s society. A libertarian state could reasonably expect the ability to make a rational assessment about proportionate punishment to be strongly correlated with the measurable aspects of rationality that it takes to be sufficient evidence of the normative power to consent to the authority of the state. The state must draw a line somewhere; if the line is defensible on its own terms, then despite the elements of arbitrariness in its construction, those on the “may not punish” side of the line are not wronged by it.

If these Lockean views, or some modernized version of them, turn out to be defensible, then Hassoun will be unable to support premise 2 of her Legitimacy Argument. It is clear that a libertarian state would sometimes have to use coercion against merely potentially autonomous persons. However, in doing so, this state could defensibly claim that it would not violate or

3. Recent scholarship on Locke reaches conclusions about what the natural right to punishment requires that are consistent with the claims I make in this paragraph. See, for example, Ward, 2009: 233: “The problems Locke identified in the state of nature are inextricably connected to the natural power to punish, which places a heavy cognitive burden on the private judgment of individuals who are expected to resist the impulses of excessive self-love and perform *ex tempore* highly complex moral reasoning related to difficult questions about reparation, restraint, deterrence, and mutual assistance.”

take away any of their rights. In dealing with merely potentially autonomous persons living on its territory, the libertarian state would not take away their rights to person or to self-defense. It could take some of their property to pay for its protective services, but this can be given an acceptable moral justification; and the state would not invade their property rights in any other problematic way. If the merely potentially autonomous persons did have a right to punish, the libertarian state would have to take it away, and this would be difficult to justify; but since, at least in general and in typical cases, they have no such right, there is no charge to answer.

However, if the libertarian state would not, in coercing the merely potentially autonomous, ever have to violate their rights, then it would not need to obtain their consent. And if it would not need to obtain their consent, it would not need to provide them with the goods and services they would need to become autonomous. Hassoun has good reason to claim that libertarian theory implies that a libertarian state would need to get the consent of all autonomous persons subject to its laws; but she cannot claim the same thing about merely potentially autonomous persons. So premise 2 does not follow from actual consent theory; and no other justification for premise 2 has been provided.⁴

3. SECOND REPLY: EXPULSION

Suppose that I am wrong about this issue, and that Hassoun can find a way to show convincingly either that there are merely potentially autonomous persons who nevertheless retain a right to punish, or that the libertarian state would, for some other reason, be morally required to get the consent of the merely potentially autonomous. Libertarians will still have another way to reject premise 2, and thereby, to resist the Legitimacy Argument. This second strategy may not be very appealing, but it strikes me as being in accord with the basic normative logic of the overall libertarian position. To understand it, let's begin to think through some details of the kind of situation Hassoun's argument must be invoking.

Of the potentially autonomous persons living in a libertarian state, some will have family members who love them and have the desire and capacity to provide them with what they need to become fully autonomous. Since their

4. Note that a non-libertarian state or international institution would coerce its citizens in many more ways, and for many more reasons, than a libertarian state would. Therefore the reply I have offered on behalf of the libertarians wouldn't be available to defenders of such a state. As a result, the objection I have been pressing does not cast doubt on the validity of Hassoun's Autonomy Argument (45) as applied to coercive international institutions designed on a non-libertarian basis.

needs are being met, they pose no special moral problem for the libertarian state. Other potentially autonomous persons will have resources of their own, perhaps obtained through inheritance or through their own labor, with which they can purchase what they need in order to become fully autonomous. Again, they pose no special moral problem. So the people we need to concern ourselves with are obviously those who do not have, and cannot obtain through legal, private voluntary transactions, the resources necessary to meet their needs. Let's say, then, that Annie is a young and impoverished orphan, trying to survive in a libertarian utopia.

Almost all of the land in the libertarian society will be privately owned. The only exceptions would be facilities associated with the state's legitimate protective role, such as police stations, courts, and military bases. It would not be difficult to justify refusing to let Annie stay in these facilities. If the society's military bases and police stations were transformed into homeless shelters, those structures would no longer be able to carry out their rights-protective functions effectively. So libertarians would be on firm ground in claiming that, if Annie is going to live in the libertarian state, she'll have to stay on private property.

Would it be permissible, though, for a property owner to allow Annie to stay on his land, but without providing her with what she needs in order to grow up and become autonomous? Before considering Hassoun's arguments, we might have thought so. But if we adopt Hassoun's view, it will now turn out that by giving Annie permission to stay on his land, the property owner in question is consenting to a situation whose moral result will be the existence of an obligation, binding on his fellow libertarian citizens, to provide Annie with resources. This is because, once the property owner gives Annie permission to stay on his land, she will be subject to the jurisdiction of the libertarian state; as a result, according to Hassoun, she will have positive welfare rights, which the property owner's neighbors will have to pay for. In allowing Annie to stay on his land, and thereby creating a situation in which his fellow citizens acquire positive duties, the property owner is arguably violating their negative rights by imposing costs on them without their permission.

Given that allowing Annie to stay would violate other citizens' rights, it would be legitimate for the citizens of the libertarian state to make an agreement, perhaps at the constitutional convention that establishes their form of government, restricting the ability of landowners to harbor indigent potentially autonomous persons such as Annie. According to the agreement I am imagining, the property owner is free to let Annie stay, but only if he also accepts an individual obligation to ensure that Annie gets what she needs to become autonomous. Obviously it would also be permissible for charitable organizations to take on such responsibilities.

It might turn out that Annie, and all those similarly situated, would end up getting their needs met due to the voluntary charity of the libertarian state's citizens. In such a happy scenario, Hassoun would have little to criticize. Of course, the happy scenario probably would not be realized. So, if no one agrees to help Annie meet her needs, under the agreement that forms part of the constitution, no one is allowed to let Annie stay on his property. With no place in which she can legally stay, Annie would effectively be expelled from the libertarian state. The result of the agreement will be that all potentially autonomous persons who remain in the libertarian state will receive what they need to become autonomous. No residents will any longer be in a position to claim libertarian welfare rights against the state.

This reply to Hassoun's argument depends on the fact that premise 2 does not explicitly address the possibility that some persons who are currently subjects of a state will not remain so. Once we notice this possibility, we would have to consider modifying premise 2 to read

(2*) For a coercive institution to be legitimate, it must ensure that *those of its subjects who remain under its jurisdiction* secure sufficient autonomy to autonomously consent to, or dissent from, its rules.

If Annie had a right to remain in the libertarian state, this modification would be of little significance. But the libertarians could arrange their institutions in such a way that, by their own standards, there is no place in the state's territory where Annie has a moral right to be. Therefore, premise 2* does not entail the conclusion of the Legitimacy Argument.⁵

Is this solution morally acceptable? I am sure that most readers will think that it is not. The thought of wretched Annie, waiting alone in the station for the train that will take her to an unknown and precarious future, is enough to inspire pity and compassion in most people, and as a result, the response that more is owed to her than she is getting. But if you had *that* intuition, would you have been attracted to natural rights libertarianism in the first place? The idea that it is permissible for all the property-owners in the society to agree to exclude Annie from their land, as a way of avoiding the obligation to meet her needs, seems so morally objectionable as to be outrageous.

5. For this solution to be available, there must be someplace outside the libertarian state for Annie to go. That is, if libertarians adopt this solution, they must oppose a global government. Given their emphasis on the practical importance of decentralization and competition between jurisdictions, I doubt that libertarians will find this claim unpalatable. Note that, as a result, the constitutional solution defended here does not in any way undermine the appeal of the Autonomy Argument (45) as directed against non-libertarian international institutions of global scope. I do not here consider the deeply problematic situation that would arise for the libertarians if other states refused to admit Annie.

But it is objectionable in the same way, to the same extent, and for the same reasons, as the libertarian position itself.

4. CONCLUSION

The history of the twentieth century showed, to everyone willing to open their eyes and learn, that markets have many important benefits and advantages. Surely, though, the important truths in the libertarian position would be better expressed through a view that rests the justification of markets on those benefits and advantages. Surely we can recognize, and value, the enormously important good consequences of allowing individuals broad scope to make their own choices in economic matters, while also recognizing the failures and limitations of markets in some instances, and the need for prudential regulation. Surely we can recognize, and deplore, the depressing tendency for governments to make problems worse in trying to solve them, and to ignore the unintended consequences of their actions, while also recognizing that in cases of externality, asymmetric information and market failure, there is sometimes no alternative to government action. Surely a consequentialist understanding of the advantages and disadvantages of markets would be better than a libertarian one.

However convincing these claims may be, though, they do not fully address the theoretical motivations behind the natural rights libertarian position. That position must be shown to be unacceptable on its own terms. Hassoun has offered us a novel and very interesting way of trying to do so. I regret to say, however, that her argument seems unsuccessful to me. For all its promise, the Legitimacy Argument turns out to leave natural rights libertarianism just as strong as it was before.

BIBLIOGRAPHY

- Hassoun, N., 2012: *Globalization and Global Justice: Shrinking Distance, Expanding Obligations*, New York: Cambridge University Press.
- Locke, J., 1982 [1689]: *Second Treatise of Government*, ed. Richard Cox, Wheeling, IL: Harlan Davidson.
- Nozick, R., 1977: *Anarchy, State and Utopia*, Basic Books.
- Pogge, T., 2002: *World Poverty and Human Rights*, Malden, MA: Polity Press.
- Sterba, J., 2005: *The Triumph of Practice over Theory in Ethics*, New York: Oxford University Press.
- Ward, L., 2009: "Locke on Punishment, Property and Moral Knowledge", *Journal of Moral Philosophy* 6: 218-244.

Social Contract Theory in the Global Context*

PETER STONE

Trinity College Dublin

Abstract

Nicole Hassoun's *Globalization and Global Justice: Shrinking Distance, Expanding Obligations* (2012) offers a novel argument for the existence of positive rights for the world's poor, and explores institutional alternatives suitable for the realization of those rights. Hassoun's argument is contractualist (in the broad sense), and makes the existence of positive rights depend upon the conditions necessary for meaningful consent to the global order. It thus provides an interesting example of social contract theory in the global context. But Hassoun's argument relies crucially upon the ambiguous nature of the concept of consent. Drawing broadly upon the social contract theory tradition, Hassoun relies upon *actual consent theory*, *democratic theory*, and *hypothetical consent theory*. Each theoretical approach makes use of its own conception of consent. Rather than select one of these conceptions over the others, she makes use of all three. In doing so, she introduces a crucial ambiguity into the terms that, on her account, a legitimate global order must satisfy. The resolution of this ambiguity will circumscribe any effort, on the part of Hassoun or others, to specify the terms of any global social contract.

Keywords: Consent, Global Justice, Globalization, Legitimacy, Nicole Hassoun, Positive Rights, Social Contract Theory.

1. INTRODUCTION

Nicole Hassoun's *Globalization and Global Justice: Shrinking Distance, Expanding Obligations* (2012) brings social contract theory to bear on the problem of global poverty. She attempts to identify some of the terms that

* Earlier versions of the argument made here were presented at the 2013 Manchester Workshops in Political Theory and at the conference "Social Contract Theory: Past Present, and Future", University of Lisbon, May 15-17, 2014. I am grateful for the comments and suggestions received at these events. Work on this paper was supported by a grant from the Arts and Social Sciences Benefaction Fund at Trinity College Dublin.

must be satisfied by any set of principles suitable for the governance of the global arena —by any *global social contract*, if you will. Among these terms Hassoun focuses upon a number of obligations the world order incurs towards the world's poor. Or, putting things the other way around, her focus is upon the rights that a global social contract must guarantee the poor. Given the persistence and severity of global poverty, the importance of her project is hard to overstate.

In bringing the idea of a social contract to bear in the global arena, Hassoun is in very good company. Some of the most important work on global justice over the past fifty years has made similar use of social contract theory (e.g., Beitz, 1999, Rawls, 1999a). But social contract theory carries with it a large amount of philosophical baggage. This baggage includes persistent philosophical problems that have proven very difficult to solve. In this short paper, I wish to focus upon one of these problems —the problem of *consent*. The concept of consent is critical to social contract theory; it does little good to speak of a contract if one cannot also speak of people consenting to that contract in some way. But there are various ways of speaking about consent in the social contract tradition. The *concept* of consent has, within that tradition, given rise to various *conceptions* of consent.¹ Any effort to defend some vision of the social contract must specify a sense in which people can be said to consent to this vision, or else risk introducing a critical ambiguity. Such ambiguity, I shall argue, bedevils Hassoun's efforts to defend the terms she wishes to attach to a global social contract.

2. CONSENT IN THE GLOBAL CONTEXT

Globalization and Global Justice offers a novel argument for the existence of positive rights for the world's poor.² Hassoun dubs this argument the *Autonomy Argument*³, and it proceeds as follows:

- (1) Coercive institutions must be legitimate.
- (2) For a coercive institution to be legitimate it must ensure that its subjects secure sufficient autonomy to autonomously consent to, or dissent from, its rules (henceforth, *sufficient autonomy*).
- (3) Everyone, to secure this autonomy, must secure some food and water, and most require some shelter, education, health care, social support, and emotional goods.
- (4) There are many coercive international institutions.

1. On the concept/conception distinction, see Rawls (1999b, 5).

2. I examine the specifics of Hassoun's argument in more detail in Stone (Forthcoming-b).

3. Hassoun refers to this argument, with minor differences in wording, as the *legitimacy* argument elsewhere in the book (92).

- (C) So, these institutions must ensure that their subjects secure food, water, and whatever else they need for sufficient autonomy (Hassoun's emphasis; Hassoun, 2012, 45; all further references will be to this book unless otherwise indicated).

Essentially, Hassoun's argument is contractualist in nature.⁴ Coercive political institutions, whether domestic or global, require consent for legitimacy. But consent requires autonomy, and the subjects of these institutions cannot grant this consent if they lack the autonomy necessary for doing so. Coercive institutions must therefore ensure that their subjects have this autonomy, or else stop coercing them. Few non-anarchists would suggest dismantling the world's coercive institutions completely, whether they be national or global, and so those institutions must meet substantive obligations to the world's poor in order to ensure them the autonomy necessary for consent. Those poor, in turn, enjoy corresponding positive rights. In a world where the poorest of the poor suffer from terrible absolute levels of deprivation—malnutrition that may be sufficient to produce cognitive impairment, for example—these rights are likely to be substantial, and impose meaningful burdens upon the global order.

Hassoun hopes that her argument will prove convincing to many readers who might otherwise be skeptical of the existence of positive rights. She frames her argument to appeal to libertarians in particular, and devotes an entire chapter specifically to them (ch. 3). In this chapter, she usefully develops the tension between the libertarian ideas of *consensualism* (people can consent to anything) and *minimalism* (only a minimal, nightwatchman state can be a legitimate state). If an individual can consent to slavery, why can't an entire society consent to a Soviet-style command economy (97)? Following A. John Simmons (2005), Hassoun believes that the tension should be resolved in favor of consensualism. But if consent really belongs at the heart of libertarian political theory, then libertarians have good reason to accept some version of the Autonomy Argument. Hassoun's Autonomy Argument, then, moves from legitimacy to consent, from consent to autonomy, and finally from legitimacy to autonomy. Legitimacy (on the part of coercive institutions) requires consent. Consent requires autonomy; those lacking autonomy cannot consent. And therefore legitimacy requires autonomy—specifically, it requires a coercive institution to ensure autonomy on the part of its subjects. It is the first move of the argument—the move from legitimacy to consent—that is especially critical. If this move is carried out successfully—if it can be established that legitimacy requires consent—then the

4. I am using the term “contractualist” in the broad sense, and not the narrow sense given to it by Thomas Scanlon (1998). In this broad sense, the term refers to “the view that morality is based on contract or agreement” (Ashford and Mulgan, 2012).

conditions required for that consent (especially the conditions of autonomy upon which Hassoun focuses her attention) should follow straightforwardly. This first move, therefore, must be made with particular care.

Does legitimacy require consent? Answering this question requires both a conception of legitimacy and a conception of consent. On the understanding of legitimacy endorsed by Hassoun, “a coercive institution is *legitimate* if, and only if, the institution has the justification-right to use coercive force.” “An institution”, Hassoun further explains, “has a *justification-right* to make coercive rules and give coercive commands if it is morally permissible for it to do so” (Hassoun’s emphasis; 47). This justification-right is a “liberty right”, or what Hohfeld would call a *privilege* (Hohfeld, 2010, 38-50). It is permissible for a legitimate coercive institution to use coercive force—that is, nobody has a right to demand that it cease doing so—but nobody need have an obligation either to obey the institution or to help it enforce its will. Obviously, this is a very limited understanding of legitimacy, one that Hassoun distinguishes from *justified authority*. A coercive institution has the latter if and only if “individuals have a moral duty to comply with its rules” (49). This is probably closer to what many people envision by legitimacy, but Hassoun’s weaker definition of the term is all she needs for her project.

Hassoun’s understanding of legitimacy, while not uncontroversial, is quite clear and consistent. Her understanding of consent, however, is quite different. This understanding, I shall argue, is profoundly ambiguous in precisely the ways characteristic of the social contract theory tradition.

Hassoun addresses her argument to liberals, but only liberals of a contractualist bent.⁵ She is not completely consistent here. On the one hand, she admits that her arguments “will not appeal to everyone who is concerned with freedom.” On the other hand, she claims that she is addressing “those who believe everyone should have some *basic freedoms*” (emphasis in original; 12). On the one hand, her argument “is intended to appeal to everyone who takes a particularly liberal commitment to freedom seriously” (43). On the other hand, she explicitly contrasts liberalism with consequentialism, including utilitarianism (8). She seems to regard consequentialism and totalitarianism as equally alien to liberalism. (Indeed, totalitarianism and consequentialism are the only rivals to liberalism explicitly mentioned by her.) Either Hassoun denies that John Stuart Mill took “a particularly liberal commitment to freedom seriously”, or else her argument is really intended to appeal only to a particular form of liberal—essentially, a liberal who emphasizes consent as the foundation for political legitimacy.

5. I discuss further Hassoun’s understanding of liberalism in Stone (Forthcoming-a).

At times, Hassoun acknowledges that the consent-driven approach she considers is not all there is to liberalism. “Not *all* liberals” she admits, “will agree that respect for persons, autonomous or not, requires refraining from forcing them to do things to which they cannot even object. Still, there is *one* important strand of liberal thought that supports this contention” (emphasis in original; 68). But at the same time, she seems very reluctant to acknowledge any theory that does not put consent at its heart as liberal, or even sane. “What would accounts of legitimacy look like that denied that subjects must have basic freedoms under coercive institutions? On such accounts, coercive institutions could be legitimate *even though people living under these institutions could not even freely object to them*. It is hard to see how such institutions would not be totalitarian” (emphasis in original; 63). Perhaps the line between consequentialism (à la John Stuart Mill) and totalitarianism is not (in Hassoun’s eyes, at least) all that thick after all. Still, the most charitable way to read Hassoun’s argument (a few inconsistencies notwithstanding) is as an appeal to the contractualist approach to liberal thought, to the exclusion of non-contractualist approaches.⁶

Hassoun thus puts social contract theory, with its emphasis upon consent, at the heart of liberal theory (even though not all liberals are contractualists). In doing so, she acknowledges the profoundly ambiguous nature of the concept of consent that lies at the heart of social contract theory. Her acknowledgement of this point is worth quoting at some length:

At the heart of liberalism is concern for individual freedom. [A] powerful strand in liberal thought... expresses the idea that the *actual* relationship between the rulers and *each person* who is ruled must be voluntary in some way. Still, those who are concerned about individual freedom disagree about what makes this relationship voluntary. On liberal communitarian theories, for instance, this relationship is voluntary if the rulers allow or support communities of appropriate kinds that need not be explicitly consensual. Other liberal theories make consent central to legitimacy. On hypothetical consent theories, for instance, the relationship between rulers and ruled is only voluntary if (reasonable) people *would* agree to be subject to the rulers’ dictates *were* they asked. Democratic theory requires more. On democratic theory, legitimacy arises through the democratic process where the majority must actually consent to the institutions to which they are subject. Perhaps the most demanding theory of this type is actual consent theory. On actual consent theory, coercive institutions are legitimate only if they secure their subjects’ actual consent (emphasis in original; 57-58).

6. I am equating here contractualist approaches with consent-based approaches. This is not an uncontroversial move, but space prohibits any defense of it here.

Hassoun seems to believe that many (though not all) liberals want political relationships to be voluntary. Some of these liberals are social contract theorists, who may be *actual consent theorists*, *democratic theorists*, or *hypothetical consent theorists*. There are liberals, however, who are not social contract theorists and still want voluntariness to be part of political relationships. Among these are liberal communitarians. Hassoun never explains how this emphasis upon voluntariness works outside of the social contract theory tradition. Moreover, she devotes little time to liberal communitarianism (59-60), and so I take contract theory to be the focus of her attention.

Hassoun thus identifies three forms of social contract theory—actual consent theory, democratic theory, and hypothetical consent theory. Each holds that political legitimacy depends upon some form of consent. They disagree regarding just what constitutes the appropriate form of consent. Rather than select one of these three forms of theory, Hassoun intends her argument to apply to them all. But as a result of this, whenever she invokes the idea that legitimacy depends upon consent—an idea upon which her Autonomy Argument depends—her account becomes ambiguous between three very different conceptions of consent.

Take actual consent theory, for example. At times, Hassoun avails herself of the conception of consent usually employed in settings unrelated to political theory. This is the conception upon which actual consent theory relies. On this conception, to consent means “to permit, approve, or agree; comply or yield.”⁷ In the political context, a person consents to an institution when she indicates (through whatever means are appropriate) her agreement to conform to its dictates. The idea is that coercive institutions are only legitimate if their subjects agree to accept them.

Hassoun appeals to this conception of consent on many occasions. She explicitly acknowledges “an idea implicit in the social contract tradition—focusing upon what social arrangements people could freely accept” (47). For people, according to Hassoun, to consent to coercive institutions is for them “to autonomously agree to their rules” (9). People must “agree to be subject to coercive institutions” (10). Moreover, Hassoun’s understanding of *why* legitimacy requires consent fits well with this understanding of consent. Coercive institutions require consent in order to be legitimate “because their subjects have a natural right to freedom” (18).⁸ People have a right to be free,

7. <http://dictionary.reference.com/browse/consent?s=t>, last accessed October 28, 2014.

8. Hassoun follows Hart (1955) here. Technically, she argues that “as long as one does not violate another’s rights, justification is necessary to abridge one’s natural right to freedom” (52; see also 49). As stated, this isn’t a very demanding condition; it just means you need to have a good reason to coerce someone, which presumably even fascists think they have. But Hassoun clearly means that because of the natural right to freedom, justification for coercion is owed *to the people coerced*, and must satisfy them. And this is effectively actual consent theory.

in other words, unless and until they agree otherwise. This agreement can take place at a high level. Hassoun writes:

Subjects may need to be able to process some information for autonomy, but they do not have to be able to agree to every single coercive rule to which they are subject. Subjects may only need to be able to autonomously agree to the general principles underlying their coercive institutions. Alternately, consent may require only that individuals autonomously agree to the general structure of coercive rules to which they are subject, not every subsidiary rule (29).

The point remains, however, that without explicit consent to coercive institutions (at some level), these institutions cannot legitimately act.

This understanding of consent has been present since the beginnings of social contract theory. But its limitations have been understood for almost as long. For anyone adhering to this conception of consent must either admit that no state has ever been legitimate, or else stretch the conception to its breaking point. No state, or coercive institution resembling a state, has ever obtained agreement to be bound by its rules from every one of its subjects, most have obtained such agreement from at best a handful of those subject to its rules (immigrants, perhaps), and many states have never obtained such agreement from *any* of their subjects. Moreover, it is hard to see how the ideal of actual consent could even be approached in the real world; in a society with even the most minimal level of diversity, it is impossible to imagine any political arrangement that all its potential subjects would be willing to authorize. Moreover, the costs of seeking actual consent on a wide scale (identifying possible terms, running those terms by citizens, adjusting those terms in light of the number unwilling to consent, etc.) are clearly prohibitive. Contract negotiations are difficult enough with two parties, let alone millions.⁹ This point poses a problem for actual consent theory, except for philosophical anarchists willing to accept the illegitimacy of all realistic political arrangements (e.g., Wolff, 1998).

This problem pushes consent theory towards counting as “consent” actions or omissions that would not count as consent in other contexts. This approach includes Locke’s reliance upon “tacit consent” —consent not overtly expressed, but made plain through one or another forms of cooperation with the government (paying taxes, using public services, etc.). But as Locke’s critics have long pointed out, it is very difficult to define “tacit consent” in a way that does not render virtually all governments into “governments by consent.” But there is another way to identify actions as “con-

9. On the central role of transaction costs in economics, see North (1990).

sensual” without reliance upon tacit consent. One could also count the expression of political opinion —primarily through voting— as an expression of consent. This is the approach taken in what Hassoun calls “democratic theory.”

Hassoun follows the social contract tradition in this regard. She acknowledges that “Actual consent theories are implausible. Few who have considered consent theory have defended actual consent since John Locke” (95). And her response is similar to that offered by those who have tried to save the social contract approach from the damning flaws of actual consent theory. At one point, she flirts with the Lockean idea of tacit consent. “It is obviously possible”, she writes, “for an institution to... secure as much (actual or tacit) consent as possible” (57, n. 44). But more typically, she appeals to a different variant of the second conception of consent employed by social contract theorists —one considerably weaker than the one involving explicit authorization. This is the understanding of consent employed within democratic theory.

Using this second conception, Hassoun appeals to “an idea implicit in the social contract tradition, that people must have basic freedoms under coercive rule” (43). This move transforms political consent into something like a right to political speech —or perhaps a right (following the U.S. Bill of Rights) to “petition” the coercive institution “for a redress of grievances.” People require autonomy so that they can “freely dissent” (56). They need to be able to “consent to, or dissent from, the rule of their coercive institutions” (28; see also 61, 62, 63). People need to be able to “shape the nature of their relationships with the coercive institutions to which they are subject” (28). Hassoun’s reframing of consent here seems to turn it into something like a right of political consultation, of the sort Rawls requires of all “well-ordered peoples” in *The Law of Peoples* (1999a, 63). This right includes “a right to dissent from the rule of coercive institutions by conscientious objection, non-violent protest, passive resistance, and so forth” (58; see also 62).

Finally, Hassoun is not afraid to appeal to the third conception of consent —hypothetical consent— although it receives relatively little of her attention. “On hypothetical consent theory”, she writes, “legitimacy requires that coercive institutions be organized according to those principles that would be chosen in an appropriately specified original position. Reasonable people in a liberally construed original position would only agree to be subject to coercive institutions if they are able to abide by, dissent from, or consent to their rule” (61). Hassoun here follows Rawls (1999b). but any meaningful form of hypothetical consent will appear very similar to this. That is, it will involve imagined consent granted under some suitably-specified set of counterfactual conditions.

Hassoun does not recognize any problem with appealing to these different conceptions of consent because she believes that the differences between them are unimportant for the purposes of her argument. Whatever brand of contractualist you are, whatever form of consent you believe political legitimacy requires, the conclusions for her argument will be the same. But this conclusion is only tenable if one states the implications of contractualism at a very high level of abstraction. According to Hassoun, “democratic, hypothetical, and actual consent theorists have to agree to this much: Legitimacy requires that subjects be free to determine their actions and shape the nature of their relationships to coercive institutions” (59). But this claim is true only if one does not look too closely at just what it means for subjects to “shape the nature of their relationships to coercive institutions.” Daniel Dennett has observed that religious believers can only agree about the existence of God so long as they do not inquire too closely as to what they mean when they use the word “God.” Dennett explains:

For a thousand years, roughly, we’ve entertained a throng of variously deanthropomorphized, intellectualized concepts of God, all more or less peacefully coexisting in the minds of “believers.” Since everybody calls his or her version “God”, there is something “we can all agree about” —we all believe in God; we’re not atheists! But of course it doesn’t work *that* well. If Lucy believes that Rock (Hudson) is to die for, and Desi believes that Rock (music) is to die for, they really don’t agree on anything, do they? (Dennett, 2006: 209; Dennett’s emphasis.)

In a similar fashion, I fear that democratic, hypothetical, and actual consent theorists agree only that they like something called “government by consent”, and not on any meaningful understanding of what “government by consent” involves.

A close look at the requirements of consent generated by these three conceptions bears this point out, in a manner that generates serious problems for Hassoun’s argument. At one point, for example, she claims that “Political liberals almost unanimously agree, for instance, that people have a right to dissent from the rule of coercive institutions by conscientious objection, non-violent protest, passive resistance, and so forth” (58). But contrary to what Hassoun believes, it is not at all clear that all three types of consent theorists discussed by her would agree to any such thing. On the one hand, many libertarians are very attracted to the idea that consent authorizes practically anything. If someone wants to consent to slavery, then so be it.¹⁰ At

10. As noted before, Hassoun argues that libertarians should embrace *consensualism* (people can consent to anything) rather than *minimalism* (only a minimal, nightwatchman state can be a legitimate state), because the two commitments are incompatible with one another.

the same time, libertarians are usually lukewarm at best about democratic rights; being able to “shape the nature of their relationship” with government seems less important to them than being able to avoid the relationship to the greatest extent possible. And why should actual consent theorists require political arrangements to include democratic rights anyway? If people grant consent to an arrangement with such rights, fine, but they could just as easily consent to some other arrangement. (Granted, it might not be advisable for anyone to consent to arrangements without such rights, but that is of no concern to the actual consent theorist.) Indeed, the entire idea of a “right to dissent” must seem strange to an actual consent theorist. Why would one be entitled to a right to object to arrangements to which one has already granted consent?

On the other hand, those deeply concerned with rights of democratic participation (like most egalitarian liberals) have little use for actual consent.¹¹ This is why Rawls, while originally attracted by Hart’s idea of a “right to liberty”, does not ground the natural duty to promote justice in anything resembling actual consent (Rawls, 1999b, §19).¹² And this difference should not be surprising at all; the conceptions are so different that they must unsurprisingly play very different roles in any theory of government employing them. Finally, the relationship between hypothetical consent and the right to “shape one’s relationship” with one’s political institutions is rather complicated. Consider the following claim Hassoun makes about hypothetical consent theory; “Reasonable people in a liberally construed original position would only agree to be subject to coercive institutions if they are able to abide by, dissent from, or consent to their rule” (61). It may be true that people would only grant hypothetical consent to political arrangements that guaranteed them certain democratic rights, such as a right to dissent.¹³ In this regard, the democratic theorist and the hypothetical consent theorist may well reach the same conclusion, and endorse the same political system. But their *reasons* for reaching this conclusion would be very different. For the democratic theorist, it would be the democratic rights that provide

other. But Hassoun’s argument for consensualism can be used against her here. If consensualism ought to trump minimalism, as Hassoun believes it should, then shouldn’t it also trump political rights as well? One could imagine a polity in which every citizen has consented to a democratic form of government, but one could just as easily imagine a polity in which everyone has consented to an authoritarian police state. While one may be (slightly) easier to imagine than the other, both must count as legitimate in the consensualist’s eyes.

11. In Albert Hirschman’s terms, libertarians tend to focus upon rights of *exit*, while democrats are centrally concerned with ensuring *voice* (Hirschman, 1972), although as Hassoun notes libertarian notions of consent involve a bit more than simply a right of exit (98).

12. Actual consent does, for Rawls, ground *obligations* (as opposed to duties) which follow from accepting positions of responsibility in a just society (Rawls, 1999b, §18).

13. This is why most social contract theorists who believe in hypothetical consent, such as Rawls, also believe in democratic rights.

the reason for calling the system consensual; the hypothetical agreement would be irrelevant. But for the hypothetical consent theorist, it would not be the democratic rights that form the critical locus of consent. It would be the hypothetical consent itself. If people could conceivably consent to arrangements without democratic rights, then this would surely satisfy the hypothetical consent theorist. The democratic theorist and the hypothetical consent theorist thus have very different things in mind when they identify political systems as “consensual”, even if both endorse democratic institutions. When the democratic theorist and the hypothetical consent theorist say that the institutions they recommend enjoy “consent”, they simply do not mean the same thing, even if they happen to be speaking about the same institutions.

The differences between these three conceptions of consent are deep and profound, despite Hassoun’s tendency to oscillate between them. In principle, they are completely independent of each other. Indeed, they will occasionally be directly at odds with each other. This would happen, for example, if an agent explicitly agreed to be subject to a coercive institution that granted no right of dissent to its subjects, or that could not be accepted by people in the correctly-specified hypothetical scenario (e.g., Rawls’ original position). This is far from a hypothetical scenario. Libertarians, for example, regularly place little emphasis upon democratic rights; their concern lies with preventing democratic majorities from tampering upon property rights. This places actual consent at the centre of their concerns. (At the same time, the contrast Hassoun draws between consensualism and minimalism suggests this libertarian position is not completely coherent.) This is in stark contrast to democratic theorists, who wish all citizens to enjoy equal rights of democratic participation, even for those who do not place a premium upon democratic rights (like libertarians). And so not surprisingly, each conception of consent has its own defenders, with the defenders of one often stridently opposed to defenders of the other.

All of this renders Hassoun’s efforts to appeal to all three conceptions of consent at once deeply problematic. Hassoun wants to convince all liberals everywhere that they should demand positive rights for everyone (especially the poor), on the grounds that institutions cannot be grounded upon consent without the provision of such rights. But if consent can mean three different things, then Hassoun is making, not one argument, but three, each relying upon a different conception of consent. And there is no reason to assume that a political arrangement satisfying one conception of consent will satisfy the other two. But without this assumption, Hassoun will not be able to win the universal liberal assent to positive rights that she seeks.

3. CONCLUSION

Hassoun never commits to one or another conception of consent in *Globalization and Global Justice*. This is not a simple oversight on her part. The goal that Hassoun wishes to achieve in the book using the Autonomy Argument is quite clear. “This book”, she writes, “may help extend the consensus on some important obligations to the poor” (12). Hassoun returns repeatedly to this goal throughout the book (e.g., 18). Hassoun is deeply concerned about the plight of the world’s poor, particularly the “bottom billion” which lacks many of the most elementary necessities. She describes this plight with a great deal of care and sympathy. And so her concern throughout the book is to win converts to a case for recognizing the people of the world —especially the poor— as possessing certain vitally important positive rights, rights which are currently being neglected and which the global political order must acknowledge.¹⁴ In order to win as many converts as possible, Hassoun appeals to as much of the social contract tradition as she can. By starting “from an idea implicit in the social contract tradition, that people must have basic freedoms under coercive rule”, —however these basic freedoms are understood— Hassoun hopes to persuade everyone who accepts this idea “that there are positive obligations to ensure that people are capable of avoiding severe poverty” (43).

Unfortunately, Hassoun’s argumentative strategy runs afoul of the real and significant differences between the various conceptions of consent to be found in the social contract tradition. One cannot argue, as Hassoun does, that legitimacy depends upon consent without specifying a conception of consent, unless the differences between conceptions are immaterial, which is highly unlikely. There may be good reasons for social contract theorists to accept Hassoun’s argument. But before one can judge this claim, one must know which version of social contract theory one has in mind.

REFERENCES

- Ashford, E., and Mulgan, T.: “Contractualism.” *The Stanford Encyclopedia of Philosophy* (Fall 2012 Edition), E. N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2012/entries/contractualism/>.
- Beitz, C. R., 1999: *Political Theory and International Relations*, New ed., Princeton, NJ: Princeton University Press.
- Dennett, D. C., 2006: *Breaking the Spell: Religion as a Natural Phenomenon*, New York: Penguin.
- Hart, H. L. A., 1955: “Are There Any Natural Rights?”, *Philosophical Review* 64: 175-91.

14. I discuss some of the other problems posed by Hassoun’s efforts to extend the consensus in (Stone, Forthcoming-a).

- Hassoun, N., 2012: *Globalization and Global Justice: Shrinking Distance, Expanding Obligations*, New York: Cambridge University Press.
- Hirschman, A. O., 1972: *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*, Rev. Paperback Ed., Cambridge, MA: Harvard University Press.
- Hohfeld, W. N., 2010: *Fundamental Legal Conceptions as Applied in Judicial Reasoning*, Clark, NJ: The Lawbook Exchange.
- North, D. C., 1990: *Institutions, Institutional Change and Economic Performance*, New York: Cambridge University Press.
- Rawls, J., 1999a: *The Law of Peoples*, Cambridge, MA: Harvard University Press.
- 1999b: *A Theory of Justice*, Rev. ed., Cambridge, MA: Belknap Press.
- Scanlon, T. M., 1998: *What We Owe to Each Other*, Cambridge, MA: Belknap Press.
- Simmons, A. J., 2005: "Consent Theory for Libertarians", *Social Philosophy and Policy* 22: 330-56.
- Stone, P., Forthcoming-a: "The Pursuit of Consensus in Global Political Theory", *Public Affairs Quarterly*.
- Forthcoming-b: "Review of *Globalization and Global Justice: Shrinking Distance, Expanding Obligations* by Nicole Hassoun", *Ethical Theory and Moral Practice*.
- Wolff, R. P., 1998. *In Defense of Anarchism*, Rev. Ed., Berkeley: University of California Press.

Sufficiency, Equality and the Consequences of Global Coercion*

KOK-CHOR TAN

University of Pennsylvania

Abstract

In some discussions on global distributive justice, it is argued that the fact that the state exercises coercive authority over its own citizens explains why the state has egalitarian distributive obligations to its own but not to other individuals in the world at large. Two recent works make the case that the global order is indeed coercive in a morally significant way for generating certain global distributive obligations. Nicole Hassoun argues that the coercive character of the global order gives rise to global duties of humanitarian aid. Laura Valentini argues that the existence of global coercion triggers global distributive duties more demanding than mere humanitarianism, but not necessarily as demanding as cosmopolitan egalitarian duties. This review essay suggests that Hassoun's and Valentini's depictions of the global order as coercive entitle them to the stronger conclusion that there are global egalitarian duties.

Keywords: Egalitarianism, sufficientarianism, global justice, coercion, cosmopolitanism, statism.

1. INTRODUCTION

Coercion plays a prominent role in some current discussions on global justice. In these debates, the fact of coercion in the global domain, or its absence thereof, determines the kinds of global obligations we are said to have or not to have. Michael Blake, for instance, has argued that egalitarian distributive obligations take hold only among persons who need to justify to each other the coercive institutional arrangements that they are jointly supporting. According to Blake, such arrangements are justifiable only if they undertake egalitarian distribution commitments to those living under them.

* Many thanks to Nicole Hassoun, Thom Brooks, Paula Casal and two readers for their very helpful comments and suggestions. I am also additionally indebted to Paula for her numerous editorial input and suggestions on structuring this review essay.

Since, so Blake argues, the relevant kind of coercion in need of justification occurs in the state or domestic context, but not in the global arena, egalitarian obligations are state-centric and not global in scope.¹ In a similar vein, Thomas Nagel argues that egalitarian obligations are activated only among individuals who are members of a shared coercive political arrangement. Since the global order does not constitute a shared coercive political association, global egalitarian obligations have no place in the global arena.²

I will consider the above to be the standard form of what has been called “coercion-based theories” of global justice. This is not to ignore that there can be variations within the standard form. In fact, there is an important difference between Nagel’s and Blake’s positions worth noting. For Nagel it is not sufficient that one finds oneself institutionally coerced for justification to be owed to one. In addition, one must also be regarded as a co-author of the institutional system in order to enjoy the standing to demand justification for the coercion. The coercive character of political institutions presents a problem of justification to members because these institutions require an “active engagement of the will of each”; they are institutions supposedly created and imposed in their name. For Blake, the problem with coercive institutions that introduces the problem of justification is the more direct one of systematic restrictions on personal autonomy (Nagel, 2005: 129). This important difference is reflected in the way each responds to the objection that (restrictive) immigration policies of countries are coercive of outsiders wanting in. Blake’s response is that this coercion is not systematic and ongoing since it is not affected via a global institutional order but through the policies of individual states. So while immigration policies can be coercive of some people, it does not constitute coercion of the relevant kind (Blake 2001: 280). In contrast, Nagel’s response is that while outsiders are coerced by the restrictive immigration rules of particular countries, they do not have the standing to demand justification for this coercion since they (as outsiders) are not considered to be co-authors of these policies (Nagel 2005: 129-30).³ Despite this difference, however, both their arguments share the following form:

- (1) There is a standing moral duty to assist people deprived of basic needs.

1. Blake, 2001; Blake further develops this account in his recent *Justice and Foreign Policy* (2013).

2. Nagel, 2005. Blake’s and Nagel’s theses have attracted much discussion, including in the two books reviewed here. For one response, see Caney, 2008. For a discussion on coercion and its connection to distributive equality, see Sangiovanni, 2007.

3. A short way of seeing the difference between the two responses is that while Blake has to show that immigration restrictions are not systematically and legally enacted at the global level, Nagel has to show why outsiders are not properly co-authors of such policies.

- (2) Egalitarian obligations, however, take hold if and only if there is coercion of a relevant kind that needs to be justified.
- (3) The domestic order is coercive in this relevant way.
- (4) The global order is not.
- (5) Conclusion: there is a domestic egalitarian obligation but no global egalitarian obligation.

So while Nagel and Blake disagree about the conditions under which coercion becomes morally relevant (as specified under [2]), their arguments share the basic commitments that (a) coercion (under certain contexts) is necessary and sufficient for generating egalitarian obligation, and that (b) the global arena is not coercive in the relevant way. Two important points of the standard account are worth highlighting for my present purpose. One is that the claim that the global order is not a coercive order (of the right kind) is used to block attempts at extending egalitarian arguments to the global context. At issue is the specific matter of global egalitarian justice. The other is that it is not a point of contention that we have a humanitarian duty to assist people in dire need. More importantly, this humanitarian duty or duty of assistance is independent of facts of coercion. For Nagel, it is simply a “pre-political” moral right persons have to be assisted under such conditions, and for Blake it is what respect for individual autonomy requires.⁴

It is not my objective to engage with Nagel’s and Blake’s theses in this review essay.⁵ I outline them to provide a framework and context for discussing two recent books that attempt to derive obligations of global justice from claims about coercion.⁶ What is common to both these works is their belief that the global order is a coercive order in a morally relevant sense. Although this is a rich claim in itself and deserving of extended discussion, I am more directly interested in the normative conclusions about global justice that these authors draw from their claims about global coercion.⁷

2. POVERTY AND GLOBAL INEQUALITY: HUMANITARIAN AND EGALITARIAN DUTIES

To start, let me clarify two different categories of global obligations, already alluded to above, that will be relevant to the present discussion. The brief

4. See Nagel, 2005: 127, 131-32; and Blake, 2001: 258, here marking the distinction between concerns of “relative deprivation” (i.e. equality) that arise only in the context of coercion, and concerns of “absolute deprivation” that arise directly from valuing personal autonomy.

5. I attempt this in Tan, 2006.

6. The two books I will review are Nicole Hassoun’s *Globalization and Global Justice: Shrinking Distance, Expanding Obligations* (Hassoun, 2012); and Laura Valentini’s *Justice In a Globalized World: A Normative Framework* (Valentini, 2011).

7. Parts of the present discussion expand on remarks in Tan, 2013.

remarks here will be mostly familiar but I rehearse them to make more concrete the backdrop of my discussion to come.

One global obligation is the duty to alleviate poverty; the other is the duty to regulate global economic inequality. We may call the former a humanitarian duty and the latter an egalitarian distributive duty. A humanitarian duty and an egalitarian duty are distinct kinds of obligations, both in their form or structure and in their objective. A humanitarian duty, the duty to counter poverty, is “sufficientarian” in form. It is non-comparative in that its benchmark is some non-relation threshold based on, for example, personal well-being, standard of living, or access to opportunities. And its objective is that of bringing individuals up to the defined standard of sufficiency. An egalitarian duty has a different structure. It is inherently relational, meaning by this that its benchmark is comparative. How well one is faring, from an egalitarian perspective, is not determined by reference to some non-relational threshold, but by reference to how well others are doing. And the objective of an egalitarian principle is to regulate the comparative gap between the advantaged and the less advantaged according to some ideal of distribution. So while a humanitarian duty is discharged when persons achieve the target of sufficiency (however that is defined), an egalitarian duty is continuous and remains in play so long as there remain inequalities to be regulated.

The distinction between humanitarian and egalitarian duties is not therefore merely semantic but in fact reflects the different forms and substantive goals of these duties.⁸ In contemporary global justice discussions, there is a tendency to describe the latter as a duty of justice and a humanitarian duty to be different from duty of justice. Depending on how an author defines a duty of “justice”, such a characterization can often be helpful in reminding us of the different categories of these duties. So, what is more important is not how we label and classify these duties, but the awareness that these are structurally and substantively different kinds of duties (as described above). So for the present purpose, nothing turns on whether humanitarian duties are duties of *justice* or not. We can count these duties, humanitarian and egalitarian, as duties of global justice so long as their structural and substantive differences are kept in sight.

Now, taking an egalitarian duty to be conceptually distinct from a humanitarian duty does not mean that concerns of equality and poverty are independent of each other. It can well be the case that global poverty cannot

8. Thus a humanitarian duty in this context should not be wrongly equated with charity or an act of supererogation. It is a moral duty and hence in this sense not optional. What distinguishes it from duties of egalitarian justice is its form and objective. For more on the difference between sufficientarianism and egalitarianism, see Casal, 2007. For completeness, I should note that on this understanding of egalitarianism (as comparative), prioritarianism (i.e., prioritizing the needs of the worst off) is non-egalitarian. Here see Parfit, 1997.

be properly mitigated while significant global inequalities between persons or societies remain a fact. In this case, the commitment to eradicate poverty results in an instrumental commitment to limit inequalities. But it is, however, also imaginable that significant inequalities can persist between countries without anyone or any society being impoverished. Or, put another way, even when all persons in the world are lifted above the poverty line, it is still possible, and likely in practice, that significant inequalities remain.

Below, I will first consider a coercion-based theory that relies on the fact of global coercion to support a humanitarian duty to protect and provide for the basic needs of persons, and then I will look at another that makes the stronger claim that legitimizing global coercion will require an obligation more robust than a duty of humanitarianism but less demanding than an egalitarian distributive duty.

3. HUMANITARIAN DUTY

Nicole Hassoun's purpose, in her book *Globalization and Global Justice*, is to clarify the basis of the humanitarian duty to assist the global poor, and in so doing she hopes to provide a grounding of this duty that can overcome the libertarian objection that there are no positive duties to provide for people's basic needs.

Philosophers typically regard the problem of world poverty to be less of a philosophical challenge and more of a problem of the lack of political will. In a sense this is right, for few people, politicians included, will openly deny that world poverty represents a serious moral failing for humanity. It is a failure to do what many people would say is required as a matter of justice. Yet the seeming lack of political will in the global response to poverty is not entirely a non-philosophical one. This inaction has to do in large measure to disagreements about the causes of world poverty, and disagreements about the basis and therefore the content and limits of our humanitarian obligations to address a recognized global problem. Moreover, there remain philosophical hold-outs, in this case, global libertarians (as we can call them), who will deny that there is any obligation to assist the global poor. It is therefore not superfluous for philosophers to address and examine the source and content of the obligation to aid the global poor. At the very least, we can achieve greater clarity about obligations we believe we have.

Hassoun's basic claim is that since the global institutional order is coercive of virtually all individuals in the world, this institutional arrangement is illegitimate unless it also actively attends to the basic subsistence needs of individuals. For Hassoun, an institution is coercive if "individuals or groups violating its rules must be likely to face sanctions for the violation... Coercion

usually creates conditions under which the coerced have no good alternative except to do what their coercer wants them to do” (Hassoun, 2014: 50).⁹ She points out that institutions such as the WTO, NATO, and the UN create and enforce rules and arbitrate between rules, and hence are coercive on her account (ibid). These are institutions attendant to the phenomenon of economic globalization or whose significance and pervasive force are enhanced because of it.

Granting this description of the global order, why does this present a problem of justice? Why is there the need to legitimize the coercive global arrangement? That is, and this is what Hassoun means by to “legitimize”, why is it necessary that people living under this order and subject to its coercive authority can see it as justified?¹⁰ Hassoun’s fundamental normative premise is that failure to make legitimate this arrangement to people living under it would amount to an unjustifiable restriction of their autonomy. Because individuals are autonomous agents, any coercive arrangements impacting them—since they restrict their choices on pain of sanctions—that cannot be justified *to them* is illegitimate. What is needed to rescue our arrangements from this crisis of legitimacy is the consent of those subject to them. However, to ensure that individuals can consent properly to these arrangements, we must do what we can to “ensure that their subjects secure food, water, and whatever else they need for autonomy” (Hassoun, 2012: 89). In other words, the duty to provide for the needs of persons derives from the duty to protect the autonomous capacity of agents, and the duty to protect this autonomous capacity derives from the duty to ensure that persons are in a position to consent to their coercive situation. In turn, individuals must be able to so consent if our global order is to be rescued from the crisis of illegitimacy.

It seems to me that the ideal of autonomy does double-duty in Hassoun’s argument. First autonomy explains why, as a default, coercive arrangements are problematic absent consent. Coercion is problematic, unless justified, because of its restrictions on autonomy. Second, the value of autonomy identifies the preconditions for the exercise of proper consent. Persons could

9. Now the differences among coercion theories will be affected by how each theory understands “coercion”. For instance, for Blake, the normatively relevant coercion he has in mind is one that is legal, systematic, and ongoing. Thus he is able to argue, against his opponents, that immigration policies of a country are not coercive of outsiders in the relevant sense. As he puts it, there is “no ongoing coercion of the sort observed in the domestic arena in the international legal arena” (Blake, 2001: 280). But the interesting difference between Hassoun and Blake concerning the global arena is not ultimately conceptual but empirical: they disagree over facts of the global order. Hassoun thinks there are global institutions that impose systematic and ongoing restrictions on persons (Hassoun, 2012: 77ff); Blake denies this.

10. For Hassoun, “Legitimacy ... is just a justification-right to exercise coercive force” against subjects who have a “natural right to freedom” (Hassoun, 2015: 18).

not consent if they are deprived of basic needs. In sum: if the global coercive order is to be legitimate, it must enjoy the consent of those it is coercing (first autonomy argument). Yet we cannot presume consent to be forthcoming, at the very least, unless the autonomous capacity of those whose consent we are seeking is developed and protected (second autonomy argument). One can't be exercising real choice for the purpose of expressing consent unless one is an autonomous moral agent. But because there are certain preconditions for agency without which we cannot possibly speak of autonomous agency, the legitimacy of coercive institutions remains in question if these institutions are imposed against a background where the preconditions for autonomy are clearly lacking. The lack of basic subsistence is just such a case. Thus it is important for the sake of winning legitimacy for our global institutional order that we accept an obligation to address basic needs deprivation.

As I will try to argue below, it is not clear what role exactly coercion has in Hassoun's argumentative strategy. It seems to me that either Hassoun begins with a sufficiently substantive conception of autonomy in order to generate a concern with coercion of the kind she has in mind, in which case, so I will suggest, that conception of autonomy is also sufficient to directly ground a commitment to poverty alleviation (thus rendering the idea of coercion redundant); or she begins with a less robust view of autonomy, in which case it is not clear how this notion of autonomy can ground a concern with coercion of the sort that her argument needs (thus rendering the idea of coercion inadequate to her cause).

To situate Hassoun's argumentative method in the larger philosophical literature, consider Thomas Pogge's account of our responsibility to address global poverty (Pogge, 2001). For Pogge, it is the fact that the global advantaged are helping (in a variety of ways) to sustain a global economic order that is harming the poor that imposes a duty of justice on the rich to respond to the poor's plight. The duty to address poverty is, in Pogge's account, a duty based in justice to make good the harms that we the rich are inflicting or have inflicted on the poor. That is, the moral starting point for Pogge is the modest one that we have the negative duty not to do harm to others. It is our violating of this negative principle not to do harm that in turn generates our positive duties to address the plight of the global impoverished. So, the rationale for Pogge's focus on the "factual" claim that the world order is harming the poor is that he wants an as ecumenical as possible a normative starting point, one that the libertarian can also find agreeable. Any plausible moral position, Pogge believes, accepts that we have a fundamental duty not to harm others, and thus the uncontroversial auxiliary duty to make amends for any harm we have caused or are causing. Thus his argument turns on the empirical matter of whether and how the world order is indeed harm-

ing the poor, and how the global rich are implicated as class in this wrong. Consider, in contrast, a different normative starting point, say that of Henry Shue's which takes that we have as a basic moral obligation the positive duty to assist those deprived of basic needs (Shue, 1979). On this account, it is immaterial whether the deprivation confronting us is caused by us or not—the fact of its existence is *sufficient* for creating an obligation on those able to respond. My point here is not to get into the debate surrounding these two approaches, but to point out that it is significant for Pogge that he establishes the harmful or coercive character of the world order because of his modest normative presumption—that our sole responsibility to each other is to avoid interfering with one another.

Hassoun adopts a similar methodology in making the case for humanitarian duties in light of the libertarian objection. Like Pogge, she does not want to butt heads with the libertarian. Instead, she prefers to engage the libertarian on the libertarian's own terms. But unlike Pogge who finds the meeting point to be the no-harm principle, Hassoun engages the libertarian's concern about legitimacy, in particular, the legitimacy of coercive institutions.¹¹

But just as Pogge's argument has generated interesting and lively discussions about whether or not he is implicitly relying on a philosophically substantial notion of harm to advance his case (accordingly calling into question whether Pogge has in fact reduced the debate on world poverty to a simply factual one), so a parallel question arises whether Hassoun is not in fact relying on a rather substantive ideal of personal autonomy to carry her argument.¹² Hassoun's invocation of autonomy invites two possible queries: One is that her normative starting point need not be one that her opponent, the libertarian, will necessarily accept; and, second, given her substantive account of autonomy, her reliance on coercion seems normatively redundant. Let me elaborate on the latter first.

Hassoun takes it to be important that the world order be exposed as a coercive one; yet she justifies her conclusion that coercion stands in need of legitimization because its potential blow on individual autonomy (in the way I tried to explain above). Indeed, she takes what she calls "The Autonomy Argument" to be crucial to her argument. Without this understanding of autonomy, specifically the preconditions of autonomy as defined, the con-

11. That libertarians should be consent theorists is of course a point of debate in the libertarian literature. Hassoun is aware of this, and engages the discussion on consent and libertarianism to defend her consent reading (96ff).

12. For some discussions on this and other matters of Pogge's approach to global justice (which I will leave aside here), see Jagger, 2010. Hassoun herself notes that she is proposing an alternative to Pogge's account because of some difficulties surrounding counts as "harming" the poor within Pogge's theory (42-43).

clusion that legitimizing coercion requires attention to basic needs does not follow.

Here a question arises: if individual autonomy is that morally significant, why can't we just draw the conclusion that we have the obligation to ensure that persons have access to food, water and so on simply because autonomy is impeded without access to basic subsistence *without having to show that they are also being coerced*? Indeed, several philosophers have made the case for a human right to basic needs on account of individual autonomy quite independently of facts about coercion or other prevailing possible wrongs.¹³

If respect for autonomy entails obligations to help provide subsistence in this more direct way, then the fact of coercion seems normatively superfluous for Hassoun's objective since the ideal of autonomy already features in the argument. What is relevant is that the global order provides for people's basic needs because individual autonomy is offended against otherwise. Indeed, Hassoun's own explication of her thesis suggests this much. Hassoun's argument proceeds in the following two steps that can be summarized as follows (Hassoun, 2012: 89). First, autonomy means that persons have the "autonomy-based" human right to food, water and other means of subsistence they need "for sufficient autonomy". Second, "to be legitimate, *coercive* institutions must do what they can to ensure that their subjects secure food, water and whatever else they need for autonomy" (*ibid.*, my italics). But if there is an autonomy-based *human right to subsistence*, it is not clear why the presence of coercive institutions is seen as a necessary condition of the duty to provide subsistence. The fact of coercion seems normatively redundant. Anyone and any institution has the responsibility to assist those deprived, irrespective of coercion.

Indeed, in the dominant debate on global justice, the fact of coercion is often presumed to be sufficient for triggering not merely humanitarian duties but egalitarian duties. In this regard, it is helpful to recall Michael Blake's position for illustration. Blake takes respect for personal autonomy alone to be sufficient to ground the concern for people's "absolute deprivation" (Blake, 2001). That is, the respect for autonomy straightforwardly engages a commitment to provide for people's basic needs. For Blake, the fact of coercion becomes significant not when we are asking about our duties in response to absolute deprivation but when we are considering a different question: do we have the duty to respond to inequality? That is, should we attend to peo-

13. For one recent attempt, see Gilabert, 2012. Gilabert argues that the obligation to assist the global poor stems directly from a cosmopolitan humanitarian concern, an obligation we have he argues independently of facts of coercion, association and so on. Hassoun herself in her book (Chapter One) offers a human rights based defense of meeting basic needs. This thus reinforces my question: what argumentative role is coercion really playing in defense of the conclusion that there is an obligation to meet basic needs?

ple's "relative deprivation" as well? It is only with regard to this question that coercion makes a normative difference —the fact of institutional coercion for Blake, coupled with the fundamental concern for autonomy, is what generates distributive egalitarian obligations. Blake's point in his paper is that since there is institutional coercion in the domestic setting but not in the global setting, we can see how one can consistently be a domestic egalitarian and not a global egalitarian. One might push Blake on his claim there is no global systemic coercion, or even challenge him on the normative premise that coercion is the *sin qua non* of global egalitarian obligations, but none of this rejects his view that there is a non-coercion based duty to provide for basic needs in virtue of personal autonomy. Hassoun in effect seems to have (unintentionally) raised the justificatory bar for global justice for even duties of humanitarian assistance have to be premised on facts about global coercion, which someone like Blake will not deem necessary. In short, in the main discussions on global justice and coercion, the dispute is not about humanitarian assistance but the stronger claim about global egalitarian obligations. For this reason, Hassoun's method of argument —invoking coercion to ground not global egalitarian obligations but humanitarian ones— is a little disconnected from this main discussion and appears a bit like a self-imposed handicap.

But this disconnect is understandable if we keep Hassoun's main target in mind. As mentioned, her concern is the libertarian who will reject even the modest duty of humanitarian aid. Blake's project to the contrary is explicitly directed at the *liberal egalitarian* who can accept the significance and implications of the ideal of autonomy. Such claims about autonomy will not hold sway with the libertarian, Hassoun's interlocutor, who will simply resist the claim that respect for autonomy enjoins the obligation to provide for persons' basic needs. The libertarian can of course endorse the importance of individual autonomy, but she will deny that this alone generates any positive obligations to provide for the conditions of the exercise of autonomy. So Hassoun thinks she needs to introduce the issues of coercion and consent in order to extend the normative implications of the libertarian ideal of autonomy.

Confronting and responding to the global libertarian is not without use, and this is Hassoun's motivating goal. Libertarians have been rather ignored in the debate on global justice for the most part since this debate, as mentioned, has largely focused on egalitarianism rather than humanitarian aid. Hassoun's discussion reminds us that global libertarianism remains a serious philosophical position that needs to be addressed.

In this regard, however, my second observation about autonomy's role in Hassoun's argument comes in. Unfortunately, it seems that Hassoun's notion of autonomy is not one that the libertarian can endorse. As said, Hassoun

wants to make the case that libertarians, who are themselves concerned foremost about the legitimacy of authority, will have to acknowledge that the global order faces a certain legitimacy crises unless individuals under its sway are in a position to give consent, and to give consent one must be autonomous to some extent. But she adds, as noted, that respect for autonomy requires that persons' basic needs be met as a precondition for exercising autonomy.

Yet it is not clear in the end if Hassoun's invocation of autonomy is really modest (or libertarian) enough to placate the libertarian. Built into her notion of autonomous consent are precisely the preconditions for autonomous choice that I believe many libertarians will reject. Libertarians who are also consent-theorists will of course require that consent be given freely and be non-rights violating. So clearly all libertarians will agree that a verbal agreement forced out of me under torture is no consent at all. But what about consent made under unfavorable economic circumstance? Whether such a consent is or is not freely given, or whether it involves rights violation or not, will depend on some background conception of individual freedom and rights, and many libertarians will disagree with the liberal egalitarian that consent under some economic stress is not freely given or is rights violating. So while Hassoun is on track when she says that the libertarian (who is also a consent theorist) will not approve of consent given under duress, she is too optimistic in thinking that the libertarian will agree that economic deprivation per se constitutes a morally relevant kind of duress. Libertarians might concede that a famished individual has fewer options in terms of what she could consent to, compared to another in a more favorable condition, but they need not conclude that such consent is thereby void. Certainly they will resist the claim that we therefore have some duty to improve the condition of the famished just so that her contracting situation is improved. The latter entails positive rights and duties that libertarians will not sign-on to. Hassoun's conception of autonomy and the obligations that she attaches to it seems, in the end, to be characteristically liberal rather than classically libertarian. It includes positive rights that libertarians will find unacceptable. So in the end, it appears that a particular conception of autonomy, one which libertarians will find hard to endorse, bears the weight of her argument.

In sum, either we accept Hassoun's ideal of autonomy (which is really a liberal rather than a libertarian conception of autonomy) and conclude that there is a duty of humanitarian aid directly on account of what it means to take autonomy seriously, in which case coercion is dispensable to the argument; or we grant the libertarian notion of autonomy, in which case, even the fact global coercion is not sufficient to show that there is any reason to take on positive duties of aid.

But if Hassoun does not succeed in her primary task, her careful account of how the global order is coercive in an ongoing, legal and institutional way has useful implications for the global justice debate. For if coercion provides at the very least a sufficient condition (if not a necessary one) for engaging egalitarian obligations, then her depiction of the global order as coercive entails egalitarian global obligations beyond the humanitarian assistance she seeks to defend. This is not to say that Hassoun must resist this global egalitarian implication of her argument, but it is a conclusion different from what she sets out to defend.

Indeed, on the matter of global egalitarianism, one might further wonder if true consent of the sort that Hassoun demands (for the purpose of legitimizing global coercion) is realizable in the presence of significant inequality between consenting parties. On her own understanding of the conditions for the exercise of autonomy (upon which true consent is predicated), which I have suggested is a characteristically liberal rather than a libertarian ideal of autonomy, one could make the argument that excessive global inequality will compromise the quality of any consent about the global order. One need not be impoverished—being disadvantaged can be enough—in order to be made an offer that is difficult to refuse. Thus the legitimacy of global coercion, on Hassoun's own principles, might demand more than a duty of humanitarian aid.

4. BEYOND HUMANITARIANISM BUT NOT QUITE EGALITARIAN?

The second coercion-based theory I will discuss will agree with my suggestion above that legitimizing global coercive arrangements will require global duties of justice more robust than humanitarian duties. But this account will disagree that these will therefore be egalitarian duties. Instead, it attempts to show that there is a third category of obligation that is entailed by the need to justify **coercive** global arrangements, and that this duty falls in between humanitarian duties and egalitarian duties in terms of its substantive content and demandingness.

In her *Justice in a Globalized World*, Laura Valentini makes the case for this middle alternative. She calls this the “third wave” in global justice which she takes to be distinct on the one side from “statism” that supports only humanitarian duties, and on the other from “cosmopolitanism” that enjoins global egalitarian duties. On her approach, “global justice requires more than statist assistance, but less than full-blown cosmopolitan equality” (Valentini, 2011: 20).¹⁴

14. There is affinity of views here with Cohen and Sabel, 2006.

As with Hassoun, the problem of coercion is a crucial part of Valentini's account. Principles of justice are principles that establish "when coercion is justified" (Valentini, 2011: 4). But while Hassoun begins from the ideal of autonomous consent, Valentini begins from the notion of freedom as independence. For Valentini, principles of justice are those principles that specify the conditions under which coercion is acceptable, and coercion introduces this consideration because it "involves non-trivial restrictions of freedom as independence" (Valentini, 2014: 178). Since principles of justice have the function of regulating and justifying coercion, justice-considerations kick-in only among agents who are interacting or institutionally engaged with each other and thus liable to be coercing one another. The key claim in Valentini's account is that the forms of coercion that provide the circumstance of justice include not just systematic coercion (i.e., coercion due to shared institutions people are imposing on each other) but also interactional coercion (i.e., direct agent-to-agent coercion) (Valentini, 2011: 15). As she puts it, understanding coercion more broadly to cover both interactional and systematic forms better captures "the multiplicity of constraints on freedom" (Valentini, 2011: 154).

The significance of this distinction between the two kinds of coercion is that each requires different kinds of moral principles for its regulation. Systematic forms of coercion are formalized and institutionalized. They shape the moral terrain in which individuals interact with each other on terms of freedom. To make acceptable this background condition to all participants, the coercive formal system of rules must be regulated by some egalitarian principles in the spirit of ensuring that it equally respects the freedom of all participants. Interactional coercion, which is coercion at the inter-personal or agential level, demands different conditions of acceptability. Since the concern here is not with a background structure that is presumed to regard the freedom of all with equal consideration, egalitarian considerations don't arise. Rather, since individual freedom can be preserved in case of interactional coercion in other ways—such as by interpersonal principles restricting or limiting intervention and by principles barring exploitation—egalitarian principles do not get activated.

Valentini agrees with those statist who hold that principles of justice play the role of legitimizing coercion within the state. But her broader reading of the forms of coercion that matter for justice also allows her to regard the global arena as a coercive one in a normatively significant way as well, and therefore also an arena in which concerns of justice have a place. Thus she departs from statist who tend, as she correctly sees it, to limit concerns of justice to the domestic state. But because the moral condition (the forms of coercion, i.e.) in the global arena is quite different from the domestic one, the content of global principles of justice will be different from domestic

principles. Specifically while domestic principles of justice will include egalitarian obligations, global principles will not. Thus her “third wave” of global justice that lies in between an overly modest statism and a too demanding “all out” cosmopolitan egalitarianism.

We will better appreciate Valentini’s broader reading of coercion and its significance to the debate by contrasting it with Blake’s, whom she regards as a representative statist theorist. As mentioned, Blake limits egalitarian justice commitments to the state because he takes *systematic* coercion to be the normatively salient form of coercion in need of justification. Thus while the global arena, Blake can concede, is coercive, it is coercive in an interactional way and so does not present a normatively significant (for purposes of egalitarian justice) form of coercion. On this matter of the coercive character of the global order, Valentini departs from Blake for two reasons. One is that, contra Blake, she believes that the global order involves *systematic* coercion even if not in as encompassing a way as domestic institutions. She gives the example of the global economic system, which is an institutional order that systematically constraints the freedom of some people through its laws and regulations (Valentini, 2011: 193ff). Second, as already noted, she utilizes a broader understanding of the types of coercion that create conditions for justice. Unlike Blake who limits his analysis to systematic coercion, Valentini regards interactional coercion as normatively salient as well. Accordingly, since the global order (as Blake will also agree) exhibits *interactional* or agent-to-agent coercion, as when one state intervenes or imposes sanctions on other, it is coercive in a normatively relevant sense for Valentini. In short, Valentini’s disagreement with Blake on the coercive character of the global order is on one front empirical (there is in fact systematic coercion) and on the other normative (interactional coercion morally matters too from the perspective of justice). The basic implication of all this is that the global order is one wherein concerns of coercion do arise as to generate considerations of justice.

Granting Valentini her broader understanding of coercion (which is instructive and illuminating) and her description of the global order as both interactionally and systematically coercive, I confine myself to this question: why aren’t global egalitarian obligations generated as a result? What is the moral difference in the forms of coercion in the domestic order and the global plane that can account for this difference?

The basic difference as Valentini sees it is that there is an encompassing institutional order or a basic structure (in the Rawlsian sense) in the domestic case that regulates and restraints individuals’ lives in a pervasive and profound way (Rawls, 1971). Given the pervasiveness and profundity of this coercive structure, it itself needs to be justified as an entity. To render this shared system acceptable to all subject to it, society must guarantee all

members equal political rights, equal opportunities, and adequate economic rights (Valentini, 2011: 176). That is, domestic justice must include some egalitarian commitments (as noted for example by the equal opportunity commitment) if the systematic coercion that is inevitable in the state is to be acceptable to all. This is where Valentini concurs generally with the main statist accounts of domestic coercion and egalitarian justice.

On the other hand, the global order exhibits a more limited case of institutionalized coercion, for example through the global economic order mixed with more common instances of interactional coercion (as mentioned earlier) that occur outside of a legal structure. Since the systematic coercion is limited and since the main form of coercion is interactional, the moral necessity here is not so much to justify a global institutional order as such (since there isn't one) as to regulate the specific global systems (like the global economic system) and the various inter-state or personal conduct. So there must be principles of justice to render this condition acceptable to all living in it, but the level justification called for in this situation is, so to say, weaker and does not include egalitarian commitments to render it acceptable to all involved. Indeed, one might think that global egalitarianism will run counter to legitimate statist interests such as the independence and self-determination of states which are the preconditions for ensuring the freedom of citizens. So while statists are wrong to think that the global order is one in which justice has no place, cosmopolitan egalitarians are mistaken in thinking that global principles will just be domestic egalitarian principles writ large.

Valentini outlines some possible implications for global justice under this systematically coercive scenario that are plausible and sensible. They include ensuring symmetrical rules of interaction "with no profit exception for the powerful", trade rules skewed to favor the less advantaged, certain redistributive commitments, compensating poor countries for "otherwise adverse effects of liberalization" for instance, greater accountability and regulation of global economic institutions and so on (Valentini, 2011: 200-201). These are just indications of what global justice would require and not meant as a complete articulation of the principles of global justice. They can be seen as a first sketch of what global justice would require minimally under her framework. But Valentini is also firm that whatever else will be required for global justice, it will not include egalitarian commitments.

Valentini does a thorough job pointing out how the global order is coercive both interactionally and systematically even if in a circumscribed way. But what seems a bit quick is her conclusion that the character of global coercion unlike domestic coercion does not activate egalitarian principles of justice. Why should this be so? Now it might be the case that *interactional* coercion is sufficiently regulated by principles of compensation, of forbear-

ance, of non-exploitation and so on, and so there is no cause to introduce egalitarian commitments. That might very well be the case, and I will not pursue this point here. But what about the presence of global *systematic coercion* that Valentini also accepts? If systematic coercion in the domestic case generates egalitarian commitments (as she agrees), why does it not do the same in the global case?

It might be because global systematic coercion is *limited* and so is affected through specific systems or particular institutional orders and not through an all encompassing global basic institutional structure. But why would that translate into no egalitarian commitments? Why can't these particular and limited systems or arrangements be regulated by egalitarian principles? For example, why can't the global trade regime be governed by a principle that says that the gains of trade should be equally distributed among the relevant parties as a default (with specifications on when departure from this default is admissible, as when it advantages the less advantaged)? This egalitarian obligation will no doubt be specific to the order or system being regulated, but it is still a global egalitarian commitment. So the fact that the institutional site that we want to justify and regulate is limited in its purview (regulating trade but not other aspects of global relations) does not alone tell us that this site cannot be regulated on egalitarian terms. It will just mean the global trade regime as a global institutional order should be governed by egalitarian principles. Why isn't this a global egalitarian commitment? A principle of distribution can be egalitarian in pattern independently of the pervasiveness of the site that is applied to. There may be something about site-limitation of this sort that precludes egalitarian regulation but arguments must be given for this. The requirement that egalitarian justice must have an institutional site is not contradicted just because the relevant global institutions to which an egalitarian principle can apply are less pervasive and encompassing than the basic structure of domestic society.

It is not implausible that global distributive obligations will have different content from domestic egalitarian distributive principles, that global justice will not simply be domestic justice extended to the world. But, again as with differences in site, that there is a difference in content alone does not render global principles non-egalitarian. There are different ways of specifying an egalitarian distributive commitment, and a global principle even though differing in content from a domestic egalitarian one can still be egalitarian. Different egalitarian principles can establish different conditions and limitations of acceptable inequality. For example, Rawls's difference principle is one articulation of an egalitarian obligation, and G. A. Cohen's "equal access to advantage" is another (Rawls, 1971; Cohen, 1989). Thus, just as there can be different sufficientarian principles that will propose different conceptions of the threshold of entitlements or flourishing persons are entitled to,

so there can be different kinds of egalitarian distributive principles with different understandings of the limits of admissible inequalities. A global distributive principle grounded on the fundamental commitment to individual freedom from domination can, we grant, limit inequality differently than a domestic principle grounded on the same fundamental commitment given the different ways personal freedom is at risk in these settings. But if this principle is in the business of regulating inequality for the sake of ensuring freedom, it is formally an egalitarian principle.

One implication of the above is that it is important not to assume that cosmopolitan egalitarians necessarily hold that global principles must be identical to domestic egalitarian principles. Some cosmopolitans may indeed have so argued but that is not what defines their position as cosmopolitan egalitarian. Cosmopolitanism is not a thesis about the content of equality (to wit that global egalitarian principles must be extensions of domestic egalitarian principles) but a thesis about the reach or scope of egalitarian commitments. But egalitarian commitments can take different shapes, and there is nothing in the ideal that egalitarian justice has global reach that requires global principles to be replicas of domestic principles. One does not forfeit one's cosmopolitan egalitarian credentials just because one offers a global principle that specifies the limits of acceptable inequality differently from a domestic principle.

To clarify, I am not arguing that global egalitarianism is the only defensible or plausible option. And certainly I have not suggested independent reasons for why a concern with coercion should create obligations of egalitarian justice. My claim is that given Valentini's concern with coercion and her engagement with the statist for whom systematic coercion is sufficient for grounding egalitarian obligations, one would want to know more why she is able to resist egalitarian conclusions under these terms.

Let me connect these remarks to the opening comments on the difference between egalitarianism and sufficientarianism. A duty of basic humanitarianism will have a lower target whereas a more robust duty of assistance will have a more demanding requirement, but both are essentially versions of sufficientarianism. Given that so-called statist in fact have different humanitarian targets in mind when they speak of "humanitarian duty", their position as a whole is more helpfully described as a sufficientarian rather than humanitarian. One can disagree with a particular statist's account of our humanitarian duty because it is too weak (covering only basic needs), and advocate instead a more demanding threshold (ensuring, in addition to basic needs, that members of a society can support functioning institutions of their own). But this does not introduce a new category of distributive duty—it is still a duty of sufficiency albeit a more demanding one. And just as there can be more or less demanding forms of sufficientarianism, so

there can be more or less demanding kinds of egalitarian distributive justice. Just because one thinks that more economic inequality is tolerable globally than domestically (and therefore global distributive principles will have a different content from domestic ones) does not mean that one is not a global egalitarian if one is still in the business of regulating global inequality. Egalitarian principles are egalitarian because of their basic form—their comparative character and objective of regulating inequality—not because of their content or the way the limits of admissible inequality are specified.

Presenting the contrast more fundamentally in terms of sufficientarianism versus egalitarianism, the interesting question is, contra Valentini, not whether there is a third alternative but what particular versions of sufficientarianism or egalitarianism to endorse. In the end, Valentini may well be able to resist the egalitarian impulse I gesture at, but this will mean that she is opting for a stronger version of sufficientarianism than basic humanitarianism (perhaps something closer to Rawls's duty of assistance) (Rawls, 1999). Ultimately the dispute concerning global distributive justice remains a dispute between two basic forms of global obligations—egalitarianism versus sufficientarianism. Although this basic dispute is further complicated because there are different theories of sufficientarianism (some more demanding than others) and different theories of global egalitarianism (some more demanding than others), it remains essentially a disagreement between two different forms of obligations. There is no third category or third wave of global distributive justice.

5. CONCLUDING REFLECTIONS

Coercion-based theories are often invoked for the purpose of marking a morally significant difference between domestic justice and global justice in order to explain why it is that certain duties obtain domestically but not globally. These theories introduce both a normative premise, namely, that coercion is a *sine qua non* for certain obligations of justice to take hold, and an empirical premise, namely, that the morally relevant kind of coercion obtains domestically but not globally. In particular, in the current debate on global justice, these theories are invoked to explain why *egalitarian commitments* obtain domestically but not globally. On what I call the standard form of these theories, the premise that the global order is not a coercive institutional order in the relevant sense is in the service of an anti-global egalitarian conclusion.¹⁵

Against these anti-global egalitarian arguments, some commentators have questioned the normative premise and have argued that there are other

15. Again, I am referring to Blake, 2001 and Nagel, 2005.

sufficient conditions for caring about distributive equality besides the need to legitimize coercion (mitigating misfortune, for example). So even if the global order is not coercive in the right way, there could be other reasons for taking on global egalitarian commitments. Another available response is to put pressure on the factual premise, and make the case for global egalitarianism by denying the observation that the global order is not a coercive order.

In their stimulating books, Hassoun and Valentini opt largely for the second route.¹⁶ But what is interesting is that they do not go on to oppose the anti-egalitarianism of the standard accounts. They deny the empirical premise, but they arrive at conclusions that are not, so they say, global egalitarian. In Valentini's case, however, the engagement with the standard accounts is clear in one respect. She wants to reject the statism that is characteristic of the standard accounts. The question I pose above, however, is whether she succeeds also in resisting the anti-egalitarianism as shared by the standard view. Hassoun's coercion-theory fits with the standard debate less neatly. As said, the standard coercion theories almost always accept that there is a humanitarian duty in response to poverty. Their purpose is not to deny that there are such duties but to deny that there is an egalitarian duty in addition. Rejecting the empirical premise that there is no global coercion in order to derive a duty of humanitarian aid is to argue for that which the main coercive accounts already affirm. But if Hassoun's project does not engage neatly with the coercion-based literature (as represented by Blake and Nagel), she also extends the coercion theory outside its normal confines, to challenge the libertarian who will deny that there is a straightforward duty of assistance.

Yet I noted that if the global order is in fact coercive in the different ways Hassoun and Valentini say it is, then on the understanding of the standard versions of coercion theories, global *egalitarian* commitments ought to be generated. Within this debate, making the case that the global order is coercive has more significant normative implications for our global obligations than what they argued for.

BIBLIOGRAPHY

- Blake, M., 2001: "Distributive Justice, State Coercion and Autonomy", *Philosophy and Public Affairs*, 30/3: 257-96.
 — 2013: *Justice and Foreign Policy*, Oxford: Oxford University Press.

16. So, although Valentini has a normatively broader reading of coercion, she also disagrees with the empirical claim of statists that there is no significant systematic coercion in the global domain. It is this particular observation of hers about the global order that I have focused on in my discussion of her book.

- Caney, S., 2008: "Global Distributive Justice and the State", *Political Studies* 56/3: 487-518.
- Casal, P., 2007: "Why Sufficiency is Not Enough", *Ethics* 117/2: 296-326.
- Cohen, G. A., 1989: "On the Currency of Egalitarian Justice", *Ethics* 99/41: 296-326.
- Cohen, J., and Sabel, C., 2006: "Extra Rempublicam Nulla Justitia?", *Philosophy & Public Affairs*, 34: 147-75.
- Gilabert, P., 2012: *From Global Poverty to Global Equality: A Philosophical Exploration*, Oxford: Oxford University Press, 2012.
- Hassoun, N., 2012: *Globalization and Global Justice: Shrinking Distance, Expanding Obligations*, Cambridge: Cambridge University Press.
- Jagger, A. (ed.), 2010: *Thomas Pogge and His Critics*, Oxford: Polity Press.
- Nagel, T., 2005: "The Problem of Global Justice", *Philosophy and Public Affairs* 33/2: 113-47.
- Parfit, D., 1997: "Equality and Priority", *Ratio* 10: 202-21.
- Pogge, T., 2001: *World Poverty and Human Rights*, Oxford: Polity Press.
- Rawls, J., 1999: *The Law of Peoples*, Cambridge, MA.: Harvard University Press.
- : *A Theory of Justice*, Cambridge, MA.: Harvard University Press.
- Sangiovanni, A., 2007: "Global Justice, Reciprocity, and the State", *Philosophy and Public Affairs* 35/1: 3-39.
- Shue, H., 1979: *Basic Rights*, Princeton: Princeton University Press.
- Tan, K-C., 2013: "The Demands of Global Justice", *Oeconomia* 13/4: 665-679.
- 2006: "The Boundary of Justice and the Justice of Boundaries: a defense of global egalitarianism", *The Canadian Journal of Law and Jurisprudence* 29/2: 319-344.
- Valentini, L., 2011: *Justice In a Globalized World: A Normative Framework*, Oxford: Oxford University Press.

Legitimate coercion: what consent can and cannot do

NICOLE HASSOUN

Binghamton University

1. A NEW ARGUMENT FOR DUTIES TO AID THE GLOBAL POOR

Globalization and Global Justice's first chapter sketches a traditional autonomy-based argument for human rights, arguing, like others have done, that coercive institutions must ensure that their subjects secure food, water, and whatever else they need for sufficient autonomy. Unfortunately, this argument is not likely to convince skeptics about positive rights. So, the book provides a new argument for significant duties to the global poor. This symposium focuses on its most controversial claim: To be legitimate, coercive institutions must ensure their subjects secure sufficient autonomy to consent to their rule. I thank Charles Goodman, Peter Stone and Kok-Chor Tan for this opportunity to sharpen this claim further in response to their insightful comments and the LEAP editors, particularly Paula Casal, for helpful comments throughout this exchange.¹

2. UNDERSTANDING LIBERTARIANISM: REPLY TO GOODMAN

In his probing commentary, Charles Goodman denies my claim that libertarians must endorse actual consent theory. Goodman starts with an interesting question: “how, exactly, would the libertarian state use coercion against merely potentially autonomous citizens?” (Goodman, 2014: ??). He wants to know what libertarian rights states violate if they do not ensure that people secure basic capacities. Goodman does not think libertarians must accept actual consent theory to avoid violating rights.²

I believe that when states claim a monopoly on the exercise of coercive force within a traditionally defined territory without securing their rights-respecting subjects' consent, they violate individuals' basic libertarian rights

1. I am also grateful to Marcus Arvan, Thom Brooks, and Darrel Moellendorf for helpful comments on this reply.

2. Libertarians' reluctance to endorse any obligation to ensure that people secure basic capacities is part of what makes my argument interesting.

to protect their rights. Libertarians often talk about rights to person, property, and punishment. Nonetheless, I believe libertarians are also concerned with protection as they do *not require* individuals to punish rights violators. So states need justification to prohibit even potentially autonomous citizens from protecting their rights (or punishing others who pose a rights-violating threat to them). Libertarians should agree that these people must be able to consent (and must actually consent) to the state, otherwise it is illegitimate. States cannot take money to pay for protective services from rights-respecting people who have not consented to give up their property. Taking money from them would violate other rights. People may prefer to maintain their right to protect their rights or hire others to do so. So those who would institute a state must ensure that potentially autonomous people secure the capacities they need to consent (and actually consent) before doing so. Of course, in our world, states already exist. Still, to legitimately coerce people in the future (to exist as legitimate states) they must ensure that everyone secures what they need to consent (and secure consent). States do not take away individuals' right to self-defense. They wrongly limit this right. They claim a monopoly on the exercise of coercive force. *They* specify which things count as self-defense and which do not. Without consent, libertarians *should* maintain that this violates the basic libertarian right to self-defense.

Moreover, libertarians will reject Goodman's claim that we do not wrong potentially autonomous people capable of punishing without violating rights by prohibiting them from doing so. Most potentially autonomous people lack this capacity. Goodman misleadingly asserts that someone with this capacity "would be no more wronged than an unusually mature and responsible fifteen-year-old who is denied the right to drive a car" (Goodman, 2014: ??). Libertarians believe these are both grievous wrongs.

As Goodman points out, some potentially autonomous people receive the help they need from their families, friends, or benefactors, but some do not. Even some of those with resources require help. They must consent to others using their resources even for this purpose.³

Goodman wrongly suggests that the state can coerce potentially autonomous people for paternalistic reasons. If states can ensure their subjects secure the capacities they need to consent, the cost of doing so cannot justify riding rough-shod over basic libertarian rights. I allow that it may be acceptable to coerce people for their own benefit. Still, I doubt libertarians will take this line in general. We cannot take sleeping people's money and give them some benefit when we can wake them. Even if it costs something to wake them and get their consent, we must do it. So I do not see how we can "con-

3. It is not enough if people can, in some way, obtain the help they need (e.g. if they participate in the right way in markets), they must have the capacities they need for consent.

clude... that, in requiring those potentially autonomous beings on their territory who own valuable resources to contribute some of those resources to the defense of their society from violence, the libertarians would not wrong those beings” (Goodman, 2014: ??). Libertarians may also deny that every person benefits from being a part of a state.⁴

Goodman plausibly claims that, to punish others, people “need those faculties which Locke calls ‘calm reason and conscience’”, but some non-autonomous people have these abilities (Goodman, 2014: ??). Some who have these faculties cannot consent to a state. There are schizophrenics, for example, who can work, drive, and defend themselves and others, but are unable to engage in much political reasoning.

Finally, consider Goodman’s example of Annie —intended to show that even when someone does need help, it may be impermissible for private citizens in a libertarian state to provide it. In laying out the case, Goodman starts by asserting that:

almost all of the land in the libertarian society will be privately owned. The only exceptions would be facilities associated with the state’s legitimate protective role, such as police stations, courts, and military bases. It would not be difficult to justify refusing to let Annie stay in these facilities. If the society’s military bases and police stations were transformed into homeless shelters, those structures would no longer be able to carry out their rights-protective functions effectively. So libertarians would be on firm ground in claiming that, if Annie is going to live in the libertarian state, she’ll have to stay on private property (Goodman, 2014: ??).

So Goodman asks if it would be permissible:

for a property owner to allow Annie to stay on his land, but without providing her with what she needs in order to grow up and become autonomous? ... by giving Annie permission to stay on his land, the property owner in question is consenting to a situation whose moral result will be the existence of an obligation, binding on his fellow libertarian citizens, to provide Annie with resources. In allowing Annie to stay on his land ... the property owner is arguably violating their negative rights by imposing costs on them without their permission (Goodman, 2014: ??).

4. That is, it is a live question why it would be in their interest to be a part of a large protective organization —some may prefer to be in small organizations and Nozick has to bring in considerations of consequences to justify depriving them of their right to do so. I believe the anarchists rightly object to this and say consent is required for consistent adherence to basic libertarian principles. The upshot of my argument is that consent leads to welfarism as opposed to anarchism.

Consider a few reasons to worry about this case. First, this response relies upon some questionable empirical claims. If people would consent to any state at all, they might consent to giving a state more land than needed to secure some public goods —like national parks. Even if this is not the case, the libertarian state could expand to help people by purchasing land from citizens to create homeless shelters. Of course, the money to do so would presumably come from taxes. So this would be a kind of libertarian welfare state. This kind of taxation poses a threat to the coherence of libertarianism, but I am completely happy if my argument establishes that libertarianism is incoherent because it both requires and prohibits a welfare state. Moreover, “by giving Annie permission to stay on his land” the property owner may not be “consenting to a situation whose moral result will be the existence of an obligation” that amounts to violating negative rights of co-citizens (Goodman, 2014: ??); many uses of property impose costs that do not violate libertarian rights. Goodman must offer more argument to make his case. That said, the idea that we must prohibit private property owners from doing what they want with their property would only provide reason to think that libertarians face another terrible dilemma, and that we should reject their view.

Goodman’s final proposal is most promising but also unsuccessful. He says that “it would be legitimate for the citizens of the libertarian state to make an agreement, perhaps at the constitutional convention that establishes their form of government, restricting the ability of landowners to harbor indigent potentially autonomous persons such as Annie” (Goodman, 2014: ??).⁵ If no one offers voluntary aid, Goodman says “Annie would [have to] effectively be expelled from the libertarian state” (Goodman, 2014: ??). He takes this as an alternative to rejecting libertarianism. However, I take it that the distinction between libertarianism and anarchism hangs on whether or not states can maintain a monopoly on coercive force within a territory without having to cede land to the non-autonomous. Libertarians believe they can. Anarchists reject this conclusion. At least, Annie cannot be removed from land she owns within the libertarian state’s borders. People would probably also violate her basic libertarian rights if they forced her to sell her land or removed her from un-owned (state) land. If all of this land is necessary to protect negative rights, that poses yet another problem for libertarians and provides reason to reject their theory.

5. Not everyone in a libertarian state would be so uncharitable.

3. BEYOND THE SOCIAL CONTRACT: REPLY TO STONE

In his thought-provoking commentary, Peter Stone worries that many liberals will reject the claim that legitimate coercive institutions must ensure that their subjects secure sufficient autonomy. Stone thinks my argument relies on an ambiguous idea of consent. Moreover, he says some liberals deny “that people have a right to dissent from the rule of coercive institutions by conscientious objection, non-violent protest, passive resistance, and so forth’ (Stone, 2014: 58).” Stone notes, for instance, that “the democratic theorist and the hypothetical consent theorist ... have very different things in mind when they identify political systems as consensual, even if both endorse democratic institutions” (Stone, 2014: ??). For the “hypothetical consent theorist, it would not be the democratic rights that form the critical locus of consent. It would be the hypothetical consent itself” (Stone, 2014: ??).

My argument does not rely on an ambiguous idea of consent. *Pace* Stone (2014: ??), I do not claim people must actually consent to coercive institutions’ rules. Contrary to what stone claims, I believe that people must have the *capacities* to consent. I give a detailed account of these capacities on which everyone must be able to do at least some (instrumental) reasoning and planning. I then derive the premise that legitimate coercive institutions must ensure their subjects secure sufficient autonomy to consent from several contractualist (and non-contractualist) theories. I argue that on plausible hypothetical, democratic, and actual consent theories, legitimate coercive institutions must ensure that their subjects secure these *capacities*. I believe we should reject theories on which rulers legitimately coerce people who lack basic freedoms under their rule. We should, for instance, reject hypothetical consent theories on which people cannot even object to coercive rule (Stone, 2014: ??). We cannot justify existing coercive institutions by appeal to the idea that people “could conceivably consent to arrangements without democratic rights” (Stone, 2014: ??). People should at least be able to maintain basic freedoms under coercive rule. Similarly, we should reject any account of democracy that does not involve a constitution, or some other means of protecting basic capacities. Such views fail to respect individual freedom. This does not beg the question against them. It provides reason to reject implausible versions of the views once we see their shortcomings clearly.

Stone, rightly, notes that my argument can be expanded. After explaining the general argumentative strategy, I take libertarianism to be the stalking horse for liberalism. I do not engage with the details of every, or even a wide range of, communitarian, democratic or hypothetical consent theories. However, I devote a whole chapter to arguing that on actual consent theory, and libertarianism, legitimate coercive institutions must ensure

their subjects secure sufficient autonomy. Moreover, even some of those I do not address—including some consequentialists— might accept this conclusion.⁶

Both Stone and Kok-Chor Tan worry, however, that libertarians will reject the idea that consent requires any basic capacities. As Tan puts it, “many libertarians are very attracted to the idea that consent authorizes practically anything. If someone wants to consent to slavery, then so be it” (Tan, 2014: ??). Tan also worries that, on libertarianism, people need not be able to consent to coercion (Tan, 2014: ??). Moreover, Stone points out that actual consent theorists and “libertarians are usually lukewarm at best about democratic rights” (Stone, 2014: ??). He says that “if people grant consent to arrangements with such rights, fine, but they could just as easily consent to some other arrangement... Indeed, the entire idea of a “right to dissent” must seem strange to an actual consent theorist” (Stone, 2014: ??).

Even if libertarians maintain that people can consent into slavery, they should not deny that free consent requires basic capacities. At least, people should be able to object to coercive rule until, and unless, they give up their right to do so. The severely mentally disabled, young children, the comatose, and those deluded by hunger cannot enter into free contracts. For contracts to be free, people must be able to consent. Libertarians defend slavery because they want to ensure the fidelity of free contracts.⁷ Still, people need basic capacities to enter into free contracts.

4. BASIC RIGHTS: REPLY TO TAN

In his helpful essay, Tan worries that my argument is circular, starting from an “autonomy-based’ human right to food, water and other means of subsistence” (Tan, 2014: ??).⁸ He suggests this undermines the strand of my argument addressed to libertarians, in particular. Libertarians notoriously deny a right to autonomy exists.

6. Philosophers like John Stuart Mill endorse basic rights, e.g. to freedom, even if they offer an indirect or instrumental justification for them that appeals to consequentialist considerations (Mill, 1983, Ch. 5). My argument rules out views that ride rough-shod over individuals’ rights and allow that it is normally acceptable to coerce some people just for others’ benefit. My argument will address consequentialists, and others who reject natural rights, as long as they agree that it is generally necessary to justify coercion to the coerced and this justification requires that people at least have the capacities they need to object, or consent, to coercive rule.

7. Stone is right that “those deeply concerned with rights of democratic participation (like most egalitarian liberals) have little use for actual consent” (Stone, 2014: ??).

8. If that were the case, Kok-Chor Tan would be right to object that since “there is an autonomy-based human right to subsistence, it is not clear why the presence of coercive institutions is ... a necessary condition of the duty to provide subsistence” (Tan, 2014: ??).

My argument does not start by assuming a right to autonomy. Rather, I address liberals who believe that people must maintain a basic minimum of freedom under legitimate coercive rule.⁹ The key idea is that people should be free to shape their relationships with their coercive institutions.¹⁰ Different liberals understand this freedom in different ways but, I argue that, on plausible liberal theories, it entails a commitment to sufficient autonomy. Although people might not be able to do without coercive institutions, they should at least get to decide for themselves how to react to their subjection.¹¹ People should at least be able to object, or consent, to coercive rule. If people require assistance to secure the requisite autonomy, and no other agent or institution provides it, their coercive institutions must do so, on pain of illegitimacy.

It is easiest to see how my argument against libertarians does not presuppose a right to autonomy. Libertarians believe people have basic rights to person, property, and to self-defense.¹² I argue that libertarians must endorse actual consent theory because coercion constrains an individual's exercise of these basic libertarian rights (Tan, 2014: ??)¹³ Roughly, if rulers claim a monopoly on coercive force over all of their rights-respecting subjects within a traditionally defined territory without securing their consent, they violate people's rights to protect themselves.¹⁴ At least rights-respecting potentially autonomous people must consent to such coercive rules. To actually consent to coercive rules, people must be able to consent.¹⁵ Moreover, this requires sufficient autonomy (the ability to reason and plan). So, if no

9. Coercion threatens to violate individual's basic freedom (as well as their equality and autonomy). It does not always undermine freedom, equality, autonomy, or harm people, but it is certainly capable of doing so.

10. I take it we should respect everyone, free or not; people should be able to shape their relationships with their coercive institutions. This is the answer to Kok-Chor Tan's question: "why should a restriction on freedom be a problem in the first place but for the fact that persons are considered autonomous"? See Tan, 2014: ??.

11. Hassoun, 2012: 58.

12. Although I am not convinced people have such inviolable rights, in addressing the libertarian, I consider what follows from this claim. One might worry that in attempting to address all liberals, the book does not adequately address any of them, but doing so fully would have been exhausting. So the book focused on providing a detailed look at how the argument might go in the hard case of the libertarian and on illustrating the general argumentative approach. I hope that it provides fertile ground for future inquiry. Whether or not one sees the book as offering a unified argument or several distinct arguments will depend on whether one looks only at the major premises of the argument or at its sub-premises.

13. FALTA TEXTO-----

14. GGJ does not address anarchists who think that states can simply cede territory to rights-respecting people who do not consent to their rule nor those who claim international institutions' coercion can be justified without ensuring that all of their subjects secure sufficient autonomy.

15. On the view that libertarians are committed to requiring actual consent, see, e.g., Simmons 1999 and Long and Machan, 2008.

one provides these people with the assistance they need, even libertarians should agree that their coercive institutions must do so.¹⁶

Tan rightly suggests that my argument provides the missing premise in an argument for egalitarian duties. I argue that many coercive *international* institutions exist. So, if coercion grounds egalitarian duties, like Michael Blake, for example, maintains, there are *international* egalitarian duties.¹⁷ I also agree with Tan that people might not have the capacity to consent to, or reject, coercive offers in situations of extreme inequality.¹⁸ None of this, however, undermines my response to the libertarian. Moreover, *pace* Tan, I do not endorse skeptics' attempts to raise "the justificatory bar for global justice ...[or] duties of humanitarian assistance" (Tan, 2014: ??) by recognizing their existence and taking the time to see where their arguments go astray. Rather, I try to extend the consensus on these important duties.

5. CONCLUSIONS

Let me recap. First, *pace* Goodman, libertarians hold that if we prohibit people from protecting their rights without consent, we wrong those capable of protecting their rights without violating others' rights. People can only be deprived of the right to protect their rights with consent. Most libertarians do not believe we can coerce such people for paternalistic reasons. Rulers cannot rightly expel these people from, at least their own, privately held property. If we must prohibit private land owners from letting non-autonomous people onto their land, that only worsens the dilemma at the heart of libertarianism. It provides further reason to reject the view. Second, *pace* Stone, my argument does not rely upon an ambiguous understanding of consent. I argue that, whatever account of consent contractualists endorse, they should agree that legitimate coercive rule requires that people have basic capacities to consent. Moreover, my argument addresses many non-contractualists, including some consequentialists. Third, *pace* Tan and Stone, even if libertarians believe that people can freely consent to whatever contracts they like, they should agree that people need some basic reasoning and planning capacities to enter into free contracts. Finally, *pace* Tan, my argument is not circular. The idea of freedom from which it starts does not presuppose, but grounds a right to, autonomy. Libertarians, for instance, hold that people only have basic rights to person, property, and punishment. I

16. On some liberal views, autonomy is constitutive of the basic freedom at issue, but my argument does not rely on this being the case.

17. I do not, however, endorse Blake's view.

18. Elsewhere (e.g. Hassoun, 2013) I defend much more robust obligations to aid people above this threshold than in the book.

argue that coercion, in constraining the exercise of these basic rights, requires justification. On libertarianism, I suggest, states must secure consent to avoid violating rights and people require sufficient autonomy to consent. Tan and Stone are right, however, that it is possible to extend my argument in many ways. It may even ground global egalitarian obligations. Much room remains for future research.

BIBLIOGRAPHY

- Goodman, C., 2014: "Libertarian Welfare Rights: Can We Expel Them?", *Law, Ethics, and Philosophy* 2: ??-??.
- Hassoun, N., 2012: *Globalization and Global Justice: Shrinking Distance, Expanding Obligations*, Cambridge: Cambridge University Press.
- 2013: "Human Rights and the Minimally Good Life", *Res Philosophica* 90, 3: 413-438.
- 2014: "Coercion, Legitimacy, and Individual Freedom: A Reply to Sondernholm", *Journal of Philosophical Research* 39: 191-198.
- Long, R., and Machan. T., 2008: *Anarchism/Minarchism: Is a Government Part of a Free Country?*, London: Ashgate Press.
- Simmons, J., 1999: "Justification and Legitimacy", *Ethics* 109: 739-771.
- Stone, P. 2014: "Social Contract Theory in the Global Context", *Law, Ethics, and Philosophy* 2: ??-??.
- Tan, S., 2014: "Coercion and Global Obligations: A Commentary", *Law, Ethics, and Philosophy* 2: ??-??.

Two pictures of the global-justice debate: A reply to Tan*

LAURA VALENTINI

London School of Economics and Political Science

1. INTRODUCTION

Kok-Chor Tan's review essay offers an internal critique of my perspective on global justice. Tan grants my coercion-based account of the triggers of justice-obligations, but takes issue with my claim that, on that account, "global justice requires more than statist assistance, but less than full-blown cosmopolitan equality" (Valentini, 2011: 20, quoted in Tan, 2014: Ref.). In particular, Tan thinks that, at the very least, my denial that egalitarian justice applies globally is under-argued. Principles of international/global justice, he suggests, may well differ *in content* from principles of domestic justice, but this need not mean that they are not *egalitarian in form*. For example, he asks: why "can't the global trade regime be governed by a principle that says that the gains of trade should be equally distributed among the relevant parties as a default (with specifications on when departure from this default is admissible, as when it advantages the less advantaged)?" (Tan, 2014: Ref.). This principle differs in content from the principles of justice that liberals defend at the domestic level—which do not concern the gains from trade—yet it seems egalitarian in form. And on the face of it, Tan suggests, there appears to be little in my coercion-based account that rules it out as a candidate demand of global justice.

Tan's thoughtful discussion gives me a welcome opportunity to clarify a misunderstanding about my view, which I suspect drives his critique. Tan takes me to hold that "whatever [is] required for global justice, it will not include egalitarian commitments" (Tan, 2014: Ref.). But this is not what I argue in *Justice in a Globalized World*. Instead, I only deny that "the egalitarian principles liberals adopt to assess domestic distributions of liberties, opportunities, and economic goods should apply to the world at large" (Valentini: 6, emphasis removed). The expression "full-blown cosmopolitan equality" in the passage quoted by Tan does not refer to formally egalitarian principles

* Many thanks to Paula Casal, Christian List, Miriam Ronzoni, and Andrew Williams, for their comments on an earlier draft of this response.

in general, but to one specific class of such principles. Once this semantic difference in our understandings of “cosmopolitan equality” is clarified, our perceived disagreement on whether there could be principles of global justice that are egalitarian in form disappears. The semantic difference, though, reveals what is probably a deeper dispute between the two of us: a “meta-dispute” about what the debate on global justice is, or should be, about. Tan and I operate with different pictures of that debate.

To better substantiate these claims, in what follows, I will (i) characterize the controversy between cosmopolitans and statist, (ii) sketch my position in *Justice in a Globalized World* and, on this basis, (iii) outline where I suspect the disagreement between Tan and myself really lies.

2. COSMOPOLITANISM AND STATISM

In *Justice in a Globalized World*, I aim to offer an account of international justice that steers a middle course between statism and cosmopolitanism. These two perspectives, as I characterize them in the book, offer different answers to what I call “the question of extension”, namely “whether [liberal-egalitarian] principles of domestic justice should extend to the world at large” (Valentini, 2011: 5). Cosmopolitans answer the question in the affirmative, statist in the negative.¹

Before proceeding further, it is worth noting that Tan’s critical discussion of my book appears to be implicitly conducted *not* against the background of this question, but of a different one: “What socio-economic distributive principles, if any, apply beyond borders?”² This question is both narrower and broader than the one I address in the book. It is broader insofar as it does not focus exclusively on *justice*. It is narrower insofar as it concentrates on the distribution of a specific class of goods, i.e., socio-economic goods, while justice *simpliciter* ranges over other types of goods too (e.g., liberties).³

Having said that, the contrast between statism and cosmopolitanism I draw in *Justice in a Globalized World* can be recast as a disagreement about socio-economic distribution, and thereby rephrased in terms of different answers to the question implicitly underpinning Tan’s critique. To do so, let

1. (Valentini, 2011: 6-10) offers an overview of these two positions: chs. 2 and 3 discuss cosmopolitanism; chs. 4 and 5 discuss statism.

2. This emerges from the section of Tan’s review titled “Poverty and Global Inequality: Humanitarian and Egalitarian Duties.”

3. By “socio-economic goods” I mean resources broadly construed. For simplicity, I do not problematize how exactly socio-economic goods are conceptualized, namely the relevant “distribuendum.” See (Gosepath, 2011) for discussion ‘equality’ is a loaded and ‘highly contested’ concept. On account of its normally positive connotation, it has a rhetorical power rendering it suitable as a political slogan (Westen, 1990).

me begin by noting that principles concerning the distribution of socio-economic goods may differ along the following three dimensions: their *grounds* (G); the distributive *pattern* they mandate (P); and the *recipients* of the relevant distribuendum (Y). The general structure of these principles is the following: “Ground G requires socio-economic goods to be distributed in line with pattern P between recipients Y.” Different ways of specifying parameters G, P, and Y lead to substantively different socio-economic distributive principles. Statists and cosmopolitans specify the parameters differently.

First, all cosmopolitans hold that *justice* grounds global distributive obligations. Justice is a special type of moral concern, one that generates obligations correlative to rights, which are in principle rightfully enforceable. For instance, the obligation to pay the (fair) price for a good that one has purchased is a paradigmatic obligation of justice: it is correlative to the seller’s right to payment and in principle rightfully enforceable, e.g., by the state. Justice is contrasted with *humanity*. Obligations of humanity are neither correlative to rights, nor rightfully enforceable. They require us to help needy strangers, using resources to which we are *entitled* on grounds of justice (Buchanan, 1987). While both cosmopolitans and statists hold that obligations of humanity apply across borders, for *some* statists humanity exhausts our international distributive obligations. Concerns of justice simply do not arise beyond the state (Nagel, 2005).

Second, cosmopolitans hold that socio-economic goods should be distributed in accordance with *broadly egalitarian* distributive patterns. In other words, they are concerned with global relative as opposed to absolute deprivation. Statists, by contrast, are unanimously committed to *sufficientarian* distributive patterns: for them, international justice requires that all relevant recipients have “enough” socio-economic goods. Once the relevant threshold is met for all recipients, inequalities above that threshold are not morally significant.

Third, cosmopolitans hold that the recipients of global socio-economic justice are *individual human beings*. Statists, by contrast, focus on internally legitimate *political communities*.

In sum, cosmopolitans typically argue that global justice is domestic justice writ large, and requires egalitarian socio-economic distributions between individuals.⁴ This is the sense in which they answer the “question of extension” in the affirmative. Statists, instead, either argue that there is no such thing as “socio-economic justice beyond borders” (only global humanity), or require justice-based sufficientarian distribution between states.

4. In turn, this conclusion may be defended on either relational, or non-relational grounds. For the former, see e.g. (Beitz, 1999) and (Pogge, 1989). For the latter, see e.g., (Caney, 2005) and (Beitz, 1983).

Against this backdrop, in *Justice in a Globalized World*, I articulate a view that attempts to steer a coherent middle course between these two families of positions.

3. A BRIEF SKETCH OF MY VIEW

After exposing what I take to be the shortcomings of cosmopolitanism and statism (Valentini, 2011: chs. 2, 3, 4, and 5), I develop a coercion-based account of the “triggers” of principles of justice (Valentini, 2011: ch. 6). I start from the liberal premise that individuals have a right to a mutually justifiable distribution of freedom, one under which each enjoys the social conditions to lead an autonomous life. Since coercion always involves non-trivial restrictions of freedom, instances of coercion stand in need of justification: they must be shown to be consistent with a mutually justifiable distribution of freedom between the parties involved (Valentini, 2011: ch. 7). I call the principles establishing the conditions under which coercion is justified “principles of justice.”

It is important to emphasize that my notion of coercion is somewhat heterodox.⁵ I do not equate coercion with the imposition of commands backed by the threat of sanctions but, more broadly, with the imposition of non-trivial constraints on freedom. Moreover, I suggest that coercion so understood comes in two variants. It can be perpetrated either by an agent (*interactional coercion*) or caused by a system of rules supported by a large enough number of agents (*systemic coercion*). From my perspective, both interactional and systemic coercion stand in need of justification.

Equipped with this normative framework, I argue that what principles of justice apply beyond borders—at any given time—depends on the forms of coercion existing in the international arena. In today’s world, both interactional and systemic coercion are present. Instances of interactional coercion primarily involve states as collective agents (e.g., think about aggression or bilateral state relations more generally). Instances of systemic coercion involve global or near-global systems of rules such as those sustaining the global economy (trade and finance).⁶

I argue that the justification of international interactional coercion demands respect for the sovereign equality of internally legitimate (i.e., reasonably just) states. Sovereign equality, in turn, requires every state to be in a position to control their affairs without being continuously interfered

5. It thereby differs from other coercion-based accounts of the triggers of demands of justice, such as (Blake, 2001).

6. (Valentini, 2011: 193) is explicit, however, that the rules governing the global economy need not exhaust global systemic coercion.

with or subtly dominated by other states and non-state actors. For example, consider the influence exercised by the United States over Latin American countries during the Cold War; or the influence that powerful corporations exercise over weak states. While states susceptible to these forms of interference may be “formally sovereign”, they are much less sovereign, from a substantive point of view, than others (Valentini, 2011: 191).⁷

My account of the justification of global systemic coercion is more tentative—as I admit in the book, and Tan rightly notes. In particular, I argue that the rules governing global finance and trade should be compatible with a mutually justifiable distribution of freedom among those falling under their purview: states and their citizens. Having said that, I also note that “[g]iven the cultural diversity and social complexity characterizing the global economy, instead of aiming for a specific and complete account of what global socio-economic justice requires, we are [...] on firmer ground simply establishing what it must exclude” (Valentini, 2011: 200).⁸ I then go on to offer a few examples of policy reforms that would lead in this direction—including the implementation of more equitable rules in WTO settings, and the creation of institutions to combat harmful tax-competition and global financial volatility. Since I am not offering a full picture of what global justice, at the systemic level, requires, but merely pointing to practices that it must exclude, my account is rather open-ended. What I offer, as the book’s subtitle suggests, is a “normative framework” for thinking about global justice, rather than a definitive account of what global justice positively requires.

4. WHAT TAN AND I REALLY DISAGREE ABOUT

Based on this sketch of the view in my book, it should be transparent that I do not deny that the “global trade regime [should] be governed by a principle that says that the gains of trade should be equally distributed among the relevant parties as a default” (Tan, 2014: ref); or indeed that principles of justice that are *egalitarian in form* should apply beyond borders. Unlike other critics of cosmopolitan egalitarianism, I simply remain agnostic about this. I adopt a minimalist strategy suggesting only what global systemic coercion must exclude in order to avoid being unjust for sure. In light of this, I could perhaps be reasonably criticized for saying too little. But Tan thinks I say too

7. For related discussion, see (Ronzoni, 2012).

8. (Valentini, 2011: Ch. 7) draws a very similar conclusion in relation to *domestic* principles of justice concerning the distribution of income and wealth. I accept that domestic justice requires equality in the distribution of civil and political liberties, as well as opportunities, but I claim that “[w]hat economic inequalities are permissible beyond [a] basic-needs threshold is a question to be answered on a case-by-case basis, and which should be ultimately decided through the democratic decision procedures of each political community” (19 and 176-77).

much, specifically, that “whatever [...] will be required for global justice, it will not include egalitarian commitments” (Tan, 2014: Ref).

I suspect that this misunderstanding is prompted by my claim that, on my view “global justice requires more than statist assistance, but less than full-blown cosmopolitan equality” (Valentini, 2011: 20). Tan reads the emphasis of this sentence as being on “equality”, and from this infers that I am generally averse to any form of egalitarianism beyond borders. But that sentence, and the book more generally, express aversion to “full-blown *cosmopolitan* equality.” As I have explained earlier, by this I mean the view that the principles of egalitarian socio-economic justice that apply domestically should extend globally, in the world as it is today. So I agree with Tan that “egalitarian commitments can take different shapes, and there is nothing in the ideal that egalitarian justice has global reach that requires global principles to be replicas of domestic principles” (Tan, 2014: Ref.). The book does not argue against egalitarian commitments so broadly construed.

Having said that, Tan’s reading of what I say may reveal a deeper disagreement between the two of us. For Tan, “ultimately the dispute concerning global distributive justice remains a dispute between two basic forms of global obligations-egalitarianism versus sufficientarianism” (Tan, 2014: Ref). That is, Tan seems to focus exclusively—at least in his response—on disagreements concerning the *distributive pattern* mandated by principles of international political morality (probably taking “individuals” for granted as the relevant recipients). But as I have suggested above, principles of global socio-economic justice might differ along other dimensions too, specifically: *grounds and recipients*.⁹ I take these further dimensions of disagreement to be central to the cosmopolitan-statist controversy as I understand it, namely the controversy I address in the book.¹⁰ It should thus be no surprise that much of the book is concerned with those other dimensions as well (especially grounds) and thereby somewhat de-emphasizes the centrality of the “sufficiency versus equality” contrast.

9. As I said earlier, the specific nature of the “distribuendum” is also a possible, and important, locus of disagreement, which I omit here for brevity’s sake.

10. Tan might object that “[o]ne does not forfeit one’s cosmopolitan egalitarian credentials just because one offers a global principle that specifies the limits of acceptable inequality differently from a domestic principle” (Tan, 2014: Ref.). This strikes me as a matter of definition, with little substantive import. What is more, I myself acknowledge a distinction between “strong cosmopolitanism” (i.e., the direct extension of domestic egalitarian justice to the global realm), and “weak cosmopolitanism”, which “places limits on permissible global socio-economic inequalities without insisting that they should coincide with those placed on domestic ones” (Valentini, 2011: 16-17). After an extensive critique of strong cosmopolitanism—my main target—in (Valentini, 2011: Ch. 3) I provisionally suggest that a “weaker” cosmopolitan position may be more defensible (without endorsing it).

Once those other dimensions are taken into account, contrary to what Tan suggests, it is not true that “there is no third category or third wave of global distributive justice” (Tan, 2014: Ref.).¹¹ My view—whether one finds it plausible or not—can be described as occupying a middle ground between statism and cosmopolitanism. Specifically, like cosmopolitanism and some (but not all) variants of statism, my view holds that demands of socio-economic justice apply beyond borders. Unlike cosmopolitanism, however, the view also acknowledges that states—provided they are internally legitimate—are important subjects of international justice. And unlike statism, it implies that global justice demands something more than the transfer of resources between “independent” political communities aimed at meeting a given threshold of sufficiency (what I call “statist assistance”): it requires deep restructuring of the rules governing international finance and trade, so as to remove clearly unjustifiable constraints on the freedom of states and their citizens.

5. CONCLUSION

Tan and I do not disagree about the possibility that international/global socio-economic justice might include principles exhibiting an egalitarian distributive pattern. Tan’s discussion of my work has been helpful in giving me the opportunity to clarify this point. That said, Tan and I may well disagree about the relative importance of the “equality versus sufficiency” distinction in the global-justice debate, and possibly about how other parameters within that debate—e.g., grounds and recipients—should be specified.

BIBLIOGRAPHY

- Beitz, C., 1983: “Cosmopolitan Ideals and National Sentiment”, *Journal of Philosophy* 80: 591-600.
 — 1999: *Political Theory and International Relations*, Princeton, N.J.: Princeton University Press.
 Blake, M., 2001: “Distributive Justice, State Coercion, and Autonomy”, *Philosophy & Public Affairs* 30: 257-96.
 Buchanan, A., 1987: “Justice and Charity”, *Ethics* 97: 558-75.
 Caney, S., 2005: *Justice Beyond Borders*, Oxford: Oxford University Press.
 Gosepath, S.: “Equality”, *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), E. N. Zalta, (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/equality>.

11. The claim isn’t true even if it is understood as meaning that principles of global distributive justice *can* only be either egalitarian or sufficientarian. Here are two simple counterexamples: “The gains of trade should be distributed so as to maximize the sum-total of utility” and “The gains of trade should be distributed in proportion to participants’ contributions.” These distributive principles exhibit neither a sufficientarian nor an egalitarian pattern.

- Miller, D., 2007: *National Responsibility and Global Justice*, Oxford: Oxford University Press.
- Nagel, T., 2005: "The Problem of Global Justice", *Philosophy & Public Affairs* 33: 113-47.
- Pogge, T., 1989: *Realizing Rawls*, Ithaca: Cornell University Press.
- Rawls, J., 1999: *The Law of Peoples: With "The Idea of Public Reason Revisited"*, Cambridge, MA: Harvard University Press.
- Ronzoni, M., 2012: "Two Conceptions of State Sovereignty and Their Implications for Global Institutional Design", *Critical Review of International Social and Political Philosophy* 15: 573-91.
- Tan, K. C., 2014: "Sufficiency, Equality and the Consequences of Global Coercion", *Law, Ethics and Philosophy* 2: ????
- Valentini, L., 2011: *Justice in a Globalized World: A Normative Framework*. Oxford: Oxford University Press.

A Brief Rejoinder to Valentini

KOK-CHOR TAN

University of Pennsylvania

Laura Valentini's gracious but incisive response to my criticism helpfully clarifies her own position and puts into greater relief the point of contention between us. She says that my claim that she rejects global egalitarianism is a "misunderstanding" since she only rejects "full-blown cosmopolitan equality" and not global egalitarianism as such. However, the basic difference between both remains unclear.

I should first note that Valentini's basic aim in her book, namely to discover a distinctive third position, is to be commended and encouraged. In fact, bold, ambitious undertakings of this sort, of developing new positions or frameworks for assessing lingering problems, is what makes for really interesting and engaging philosophical work. So my rejoinder here should not be read as a dismissal of her project. To the contrary, I will put pressure on her claim in order to encourage more investigation into the possibility of a third position on global justice along the lines she suggests.

Her approach to global justice, she stresses, is open to the possibility that certain global practices be regulated by egalitarian principles (of the sort I hinted at) although she did not argue for them in her book. But if this is so, it only reinforces my central objection that her approach to global justice is not a distinctive position that provides a third alternative to statist humanitarianism and cosmopolitan equality. As she says, my reading of her "reveals a deeper disagreement between the two of us", and this is the significant point. This disagreement in the first instance has to do with how we characterize cosmopolitan equality and what would make a position non-cosmopolitan.

Whether we should call a position "cosmopolitan egalitarian" or not might seem to some readers to be a pedantic dispute. But in the context of the present discussion this is more than a terminological quibble, for my claim in my review is that the interesting and crucial difference between humanitarian obligations and cosmopolitan egalitarian obligations reduces to the difference between sufficiency and equality (or sufficientarianism and egalitarianism). What is substantively at stake in the debate on global distributive justice, in fact what I would consider the fundamental challenge, is whether global distributive principles should be sufficientarian or egalitar-

ian. Valentini, in contrast, takes the dispute to involve not just the pattern of distribution but also the grounds and recipients of distribution. Our disagreement, as Valentini also points out, far from being merely semantic, is ultimately a disagreement about what is really at issue in the debate.

In my view, Valentini's position as clarified in her reply is more clearly a variant of cosmopolitan egalitarianism. Valentini offers two reasons why her position is distinctive from cosmopolitan egalitarianism and occupies a middle position. One of them is that she takes the basic recipients of distribution to be states. This alone, however, does not make her position anti-cosmopolitan since she also conditions the normative standing of states on their respecting the equal freedom of their respective individual members (freedom here defined as independence). Cosmopolitans do regard individuals, in the world as a whole, as the basic units of equal moral concern and, in this sense, adopt a normative individualist position. But this does not mean that a global distributive principle that takes states to be the recipients is non-cosmopolitan, if the qualification of states to be recipients of justice is contingent on how they treat individuals within their borders. For instance, if states' normative status as recipients of global justice is conditional on the respect for the equal freedom of each individual (as in Valentini's theory), this is not fundamentally distinguishable from the cosmopolitan position. Cosmopolitans can allow states to be the main recipients of global distribution for a variety of reasons, including administrative, heuristic, or appeal to the virtues of a division of labour. Their position remains basically cosmopolitan if what fundamentally guides their vision is how individuals fare under the distributive arrangement.

Her other reason for distinguishing her position from a cosmopolitan egalitarian one is her understanding of the grounds of global obligations. But the ground of a principle and the pattern of the principle are distinct, and a principle is egalitarian depending on its pattern and not its ground. What distinguishes cosmopolitan *egalitarianism* from statist *humanitarianism* is not that the former is grounded in justice and the latter is not, for it is open to humanitarians to say that their minimalist commitments are nonetheless obligations of justice (and some will indeed hold this). More to the present point, cosmopolitan egalitarianism need not be tied to any particular ground of equality. Valentini herself recognizes this (allowing that there can be relational and non-relational grounds for cosmopolitan equality), but this then only further confirms my point that the ground of a principle is distinct from its pattern.

To make my central point from a different direction: Valentini means to only reject "full-blown" cosmopolitan equality, meaning by this the position that egalitarian principles for the global domain are identical to those for the domestic one. That is, on this view, global egalitarianism is just domes-

tic egalitarianism writ large. As I noted in my review, while this is certainly descriptive of some cosmopolitan egalitarian positions, “full-blown” cosmopolitan equality is only a variant of cosmopolitan egalitarianism. Valentini may well be correct that global egalitarian justice is not simply domestic justice replicated in the global arena but should take the more limited though egalitarian form she suggests. But to claim that this is a distinctive (and non-cosmopolitan) position and not just a variant of cosmopolitan equality distracts from the key dispute about global distributive justice, which is that of sufficiency versus equality. And appealing to the dimensions of grounds and recipients does not necessarily provide a resolution of this quarrel nor does it bring the two sides to a common middle ground. Statist humanitarians can accept that humanitarian obligations are obligations of justice but insist that these global obligations remain sufficientarian in form; and cosmopolitan egalitarians can accept that states are the recipients of distributive justice but insist that the global distribution remain egalitarian in form.

In short, Valentini’s thesis that there is a third approach rests on a particular characterization of cosmopolitan egalitarianism which includes features that are not essential to the view. However, since cosmopolitan *egalitarianism* is neither (i) uniquely based on some specific moral grounding nor (ii) incompatible with involving the state in its distributive enterprise, Valentini’s acceptance of global distributive principles with egalitarian pattern and whose normative starting point is the equal freedom of individuals puts her on the cosmopolitan side even though she denies the team colours. As noted above, this is not a quibble over a label but a comment on what the debate between statist humanitarianism and cosmopolitan egalitarianism is really about.

